



ДЕРЖАВНА СЛУЖБА
ІНТЕЛЕКТУАЛЬНОЇ
ВЛАСНОСТІ
УКРАЇНИ

УКРАЇНА

(19) **UA** (11) **111927** (13) **C2**
(51) МПК (2016.01)
G10L 19/00

(12) ОПИС ДО ПАТЕНТУ НА ВИНАХІД

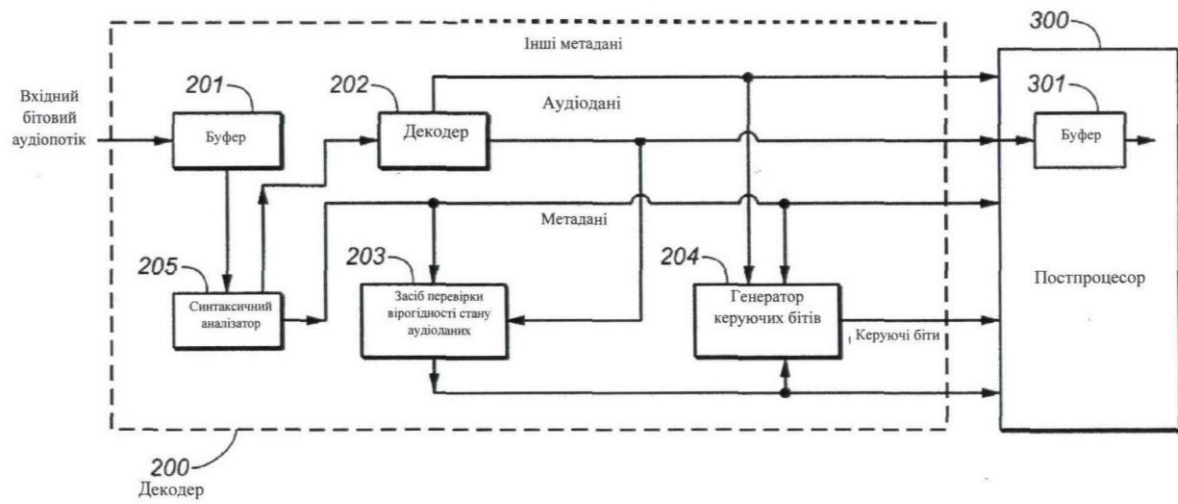
(21) Номер заявки:	а 2015 08059	(72) Винахідник(и):	Рідміллер Джеффрі (US), Вард Майкл (US)
(22) Дата подання заявки:	12.06.2014	(73) Власник(и):	ДОЛБІ ЛАБОРАТОРІС ЛАЙСЕНЗІН КОРПОРЕЙШН, 1275 Market Street, San Francisco, California 94103, USA (US)
(24) Дата, з якої є чинними права на винахід:	24.06.2016	(74) Представник:	Михайлюк Ганна Валентинівна, реєстр. №184
(31) Номер попередньої заявки відповідно до Паризької конвенції:	61/836,865	(56) Перелік документів, взятих до уваги експертизою:	US 20090097821 A1, 16.04.2009 US 20100027837 A1, 04.02.2010 US 8285791 B2, 09.10.2012
(32) Дата подання попередньої заявки відповідно до Паризької конвенції:	19.06.2013		
(33) Код держави-учасниці Паризької конвенції, до якої подано попередню заявку:	US		
(41) Публікація відомостей про заявку:	10.12.2015, Бюл.№ 23		
(46) Публікація відомостей про видачу патенту:	24.06.2016, Бюл.№ 12		
(86) Номер та дата подання міжнародної заявки, поданої відповідно до Договору РСТ	PCT/US2014/042168, 12.06.2014		

(54) АУДИОКОДЕР І АУДИОДЕКОДЕР З МЕТАДАНИМИ ВІДОМОСТЕЙ ПРО ПРОГРАМУ АБО СТРУКТУРИ ВКЛАДЕНИХ ПОТОКІВ

(57) Реферат:

Пристрій і способи генерування кодованого бітового аудіопотоку полягає у включенні в бітовий потік метаданих структури вкладених потоків (SSM) і/або метаданих відомостей про програму (PIM) і аудіоданих. Інші особливості являють собою пристрій і способи декодування такого бітового потоку й модуль обробки аудіоданих (наприклад кодер, декодер або постпроцесор), виконаний (наприклад запрограмований) з можливістю виконання якого-небудь із варіантів здійснення зазначеного способу й утримуючий буферну пам'ять, що зберігає щонайменше один кадр бітового аудіопотоку, згенерований відповідно до будь-якого із варіантів здійснення зазначеного способу.

UA 111927 C2



Фіг.3

Перехресне посилання на споріднені заявки

Дана заявка заявляє пріоритет попередньої заявки на патент США №61/836865, поданої 19 червня 2013 р., яка посиланням повністю включається в цей документ.

Галузь технічного застосування

Винахід відноситься до обробки звукових сигналів і, зокрема, до кодування й декодування бітових потоків аудіоданих з метаданими, що служать ознаками структури вкладених потоків і/або відомостей про програму відносно звукового вмісту, що вказується цими бітовими потоками. Деякі варіанти здійснення винаходу генерують або декодують аудіодані в одному з форматів, відомих як Dolby Digital (AC-3), Dolby Digital Plus (Enhanced AC-3, або E-AC-3) або Dolby E.

Передумови винаходу

Dolby, Dolby Digital, Dolby Digital Plus і Dolby E є торговельними марками Dolby Laboratories Licensing Corporation. Dolby Laboratories представляє власні реалізації AC-3 і E-AC-3, відомі, відповідно, як Dolby Digital і Dolby Digital Plus.

Модулі обробки аудіоданих, як правило, діють наосліп і не приділяють увагу історії обробки аудіоданих, що відбувалася перед прийманням цих даних. Це може працювати в інфраструктурі обробки даних, де всю обробку й кодування аудіоданих для різноманітних цільових пристроїв представлення мультимедійних даних здійснює єдиний суб'єкт, у той час як цільовий пристрій представлення мультимедійних даних здійснює все декодування й представлення цих кодованих аудіоданих. Однак така обробка даних наосліп не дуже добре підходить (або зовсім не підходить) для ситуацій, у яких безліч модулів обробки аудіоданих розкидані по різноманітній мережі або розміщені послідовно (тобто в ланцюгу) і, як очікується, оптимально виконують відповідні їм типи обробки аудіоданих. Наприклад, деякі аудіодані можуть бути закодовані для високопродуктивних мультимедійних систем, і на всьому протязі ланцюга обробки мультимедійних даних може виникнути необхідність у їх перетворенні в наведену форму, що підходить для мобільного пристрою. Відповідно, модуль обробки аудіоданих може без необхідності виконувати обробку аудіоданих одного з типів, яка вже була виконана. Наприклад, модуль регулювання рівня гучності може виконувати обробку вхідного аудіокліпу незалежно від того, чи було таке ж або аналогічне регулювання рівня гучності виконане раніше на цьому вхідному аудіокліпі. У результаті модуль регулювання рівня гучності може виконувати регулювання рівня навіть тоді, коли це не є необхідним. Така обробка даних, що не є необхідною, також може викликати погіршення й/або усунення характерних ознак під час представлення вмісту аудіоданих.

Стислий опис винаходу

В одному з класів варіантів здійснення винахід являє собою модуль обробки аудіоданих, здатний декодувати кодований бітовий потік, що містить метадані структури вкладених потоків і/або метадані відомостей про програму (а також, необов'язково, і інші метадані, наприклад, метадані стану обробки гучності) щонайменше в одному сегменті, щонайменше одного кадра бітового потоку, і аудіодані щонайменше ще в одному сегменті цього кадра. У цьому документі термін «метадані структури вкладених потоків» (або «SSM») означає метадані кодованого бітового потоку (або набору кодованих бітових потоків) структури, що служать ознакою, вкладених потоків звукового вмісту цього кодованого бітового потоку (потоків), а термін «метадані відомостей про програму» (або «PIM») означає метадані кодованого бітового потоку, що служать ознакою, щонайменше, однієї звукової програми (наприклад, двох або більшої кількості звукових програм), де метадані відомостей про програму служать ознакою, щонайменше, однієї властивості, або характеристики, звукового вмісту, щонайменше, однієї зазначеної програми (наприклад, метадані, що вказують тип або параметр обробки даних, виконаної на аудіоданих цієї програми, або метадані, що вказують, які канали програми є активними каналами).

У типових випадках (наприклад, тоді, коли кодований бітовий потік являє собою бітовий потік AC-3 або E-AC-3), метадані відомостей про програму (PIM) служать ознакою відомостей про програму, які практично неможливо перенести в інших частинах бітового потоку. Наприклад, PIM можуть служити ознакою обробки даних, застосованої до аудіоданих PCM перед кодуванням (наприклад, кодуванням AC-3 або E-AC-3), коли смуги частот звукової програми закодовані з використанням спеціальних методик кодування звуку, і профілю стиснення, що використовується для створення даних стиснення динамічного діапазону (DRC) у цьому бітовому потоці.

В іншому класі варіантів здійснення спосіб включає етап ущільнення кодованих аудіоданих з SSM і/або PIM у кожному кадрі (або кожному з щонайменше деяких кадрів) бітового потоку. При типовому декодуванні декодер витягає SSM і/або PIM з бітового потоку (що включає

синтаксичний аналіз і розуцільнення SSM і/або PIM і аудіоданих) і обробляє аудіодані для генерування потоку декодованих аудіоданих (і, у деяких випадках, також виконує адаптивну обробку цих аудіоданих). У деяких варіантах здійснення декодовані аудіодані й SSM і/або PIM направляються з декодера в постпроцесор, виконаний з можливістю адаптивної обробки даних на декодованих аудіоданих з використанням SSM і/або PIM.

В одному із класів варіантів здійснення способу кодування за винаходом генерує кодований бітовий аудіопотік (наприклад, бітовий потік AC-3 або E-AC-3), що містить сегменти аудіоданих (наприклад, сегменти AB0-AB5 кадра, показаного на Фіг. 4, або всі або деякі із сегментів AB0-AB5 кадра, показаного на Фіг. 7), що містять кодовані аудіодані, і сегменти метаданих (що містять SSM і/або PIM а також, необов'язково, інші метадані), ущільнені з тимчасовим поділом із сегментами аудіоданих. У деяких варіантах здійснення кожний сегмент метаданих (іноді іменований у цьому документі «контейнером») має формат, що містить заголовок сегмента метаданих (а також, необов'язково, інші обов'язкові, або «базові», елементи) і одне або кілька корисних завантажень метаданих, що слідує за заголовком сегмента метаданих. Метадані SIM, якщо вони присутні, укладені в одному з корисних завантажень метаданих (що ідентифікуються за допомогою заголовка корисного завантаження й, як правило, що мають формат першого типу). Метадані PIM, якщо вони присутні, містяться в іншому корисному завантаженні метаданих (що ідентифікуються за допомогою заголовка корисного завантаження й, як правило, що має формат другого типу). Аналогічно, інші типи метаданих (якщо вони присутні) містяться в інших корисних завантаженнях метаданих (що ідентифікуються за допомогою заголовка корисного завантаження й, як правило, що мають формат, специфічний для цього типу метаданих). Цей ілюстративний формат уможливорює зручний доступ до SSM, PIM і інших метаданих в інші моменти часу, ніж під час декодування (наприклад, доступ постпроцесора слідом за декодуванням, або для процесора, виконаного з можливістю розпізнавання метаданих без виконання повного декодування на кодованому бітовому потоці), і уможливорює зручне й ефективно виявлення й виправлення помилок (наприклад, помилок ідентифікації вкладених потоків) у ході декодування бітового потоку. Наприклад, під час відсутності доступу до SSM в ілюстративному форматі декодер може невірно ідентифікувати правильну кількість вкладених потоків, асоційованих із програмою. Одне корисне навантаження метаданих у сегменті метаданих може містити SSM, друге корисне навантаження метаданих у сегменті метаданих може містити PIM, а також, необов'язково, щонайменше, ще одне корисне навантаження метаданих у сегменті метаданих може містити інші метадані (наприклад, метадані стану обробки гучності, або «LPSM»).

Стислий опис графічних матеріалів

Фіг. 1 - блок-схема одного з варіантів здійснення системи, яка може бути виконана з можливістю виконання одного з варіантів здійснення способу за винаходом.

Фіг. 2 - блок-схема кодера, що представляє собою один з варіантів здійснення модуля обробки аудіоданих за винаходом.

Фіг. 3 - блок-схема декодера, що представляє собою один з варіантів здійснення модуля обробки аудіоданих за винаходом, і пов'язаного з ним постпроцесора, що представляє собою ще один варіант здійснення модуля обробки аудіоданих за винаходом.

Фіг. 4 - схема кадра AC-3, що містить сегменти, на які він розділений.

Фіг. 5 - схема сегмента відомостей про синхронізацію (SI) кадра AC-3, що містить сегменти, на які він розділений.

Фіг. 6 - схема сегмента відомостей про бітовий потік (BSI) кадра AC-3, що містить сегменти, на які він розділений.

Фіг. 7 - схема кадра E-AC-3, що містить сегменти, на які він розділений.

Фіг. 8 - схема сегмента метаданих кодованого бітового потоку, згенерованого відповідно до одного з варіантів здійснення винаходу, в тому числі заголовка сегмента метаданих, який містить синхрослово контейнера (що ідентифікується на Фіг. 8 як «container sync») і значення версії й ідентифікатора ключа (key ID), за якими ідуть корисні навантаження метаданих і біти захисту.

Позначення й термінологія

Усюди в даному розкритті, включаючи формулу винаходу, вираження виконання операції «на» сигналі або даних (наприклад, фільтрації, масштабування, перетворення або застосування коефіцієнта підсилення до сигналу або даних) використовується в широкому змісті для позначення виконання операції безпосередньо на сигналі або даних, або на обробленій версії цього сигналу або даних (наприклад, на версії сигналу, що перетерпів попередню фільтрацію або попередню обробку даних перед виконанням на ньому цієї операції).

Усюди в даному розкритті, включаючи формулу винаходу, вираз «система»

використовується в широкому змісті для позначення пристрою, системи або підсистеми. Наприклад, підсистема, що реалізує декодер, може йменуватися системою декодера, а система, що містить таку підсистему (наприклад, система, що генерує X вихідних сигналів у відповідь на кілька введень, причому ця підсистема генерує M з уведень, а інші X-M уведення

5 приймаються із зовнішнього джерела), також може йменуватися системою декодера.

Усюди в даному розкритті, включаючи формулу винаходу, термін «процесор» використовується в широкому змісті для позначення системи або пристрою, що можуть бути запрограмовані або іншим чином сконфігуровані (наприклад, програмним або програмно-апаратним забезпеченням) для виконання операцій на даних (наприклад, на аудіоданих, відеоданих або даних інших зображень). Приклади процесорів включають вентиляну матрицю з експлуатаційним програмуванням (або іншу інтегральну мікросхему, або набір мікросхем що можуть бути сконфігуровані.), процесор цифрової обробки сигналів, що може бути запрограмований або іншим чином сконфігурований для виконання конвеєрної обробки аудіоданих або інших звукових даних, процесор загального призначення або комп'ютер, що

15 може бути запрограмований, і корпусований мікропроцесор або набір мікросхем, що можуть бути запрограмовані.

Усюди в даному розкритті, включаючи формулу винаходу, вирази «процесор аудіоданих» і «модуль обробки аудіоданих» використовуються взаємозамінним чином й у широкому змісті для позначення системи, виконаної з можливістю обробки аудіоданих. Приклади модулів обробки аудіоданих включають, без обмеження, кодери (наприклад, перетворювачі коду), декодери, кодеки, системи попередньої обробки даних, системи заключної обробки даних і системи обробки бітових потоків (іноді іменовані інструментальними засобами обробки бітових потоків).

20 Усюди в даному розкритті, включаючи формулу винаходу, вираз «метадані» (кодованого бітового потоку) відноситься до окремих даних, що відрізняються від відповідних аудіоданих бітового потоку.

Усюди в даному розкритті, включаючи формулу винаходу, вираз «метадані структури вкладених потоків» (або «SSM») означає метадані кодованого бітового аудіопотоку (або набору кодованих бітових аудіопотоків) структури, що служать ознакою вкладених потоків звукового

30 вмісту кодованого бітового потоку (потоків).

Усюди в даному розкритті, включаючи формулу винаходу, вираз «метадані відомостей про програму» (або «PIM») означає метадані кодованого бітового аудіопотоку, служать ознакою, щонайменше, однієї звукової програми (наприклад, двох або більшої кількості звукових програм), де зазначені метадані служать ознакою, щонайменше, однієї властивості, або характеристики, звукового вмісту, щонайменше, однієї зазначеної програми (наприклад, метадані, що вказують тип або параметр обробки даних, виконаної на аудіоданих цієї програми, або метадані, що вказують, які канали цієї програми є активними каналами).

Усюди в даному розкритті, включаючи формулу винаходу, вираз «метадані стану обробки даних» (як, наприклад, у виразі «метадані стану обробки гучності») відноситься до метаданих (кодованого бітового аудіопотоку), асоційованих з аудіоданими цього бітового потоку, що вказують стан обробки відповідних (асоційованих) аудіоданих (наприклад, обробка даних якого типу (типів) вже була виконана на цих аудіоданих), і, як правило, також вказують щонайменше одну ознаку, або характеристику, цих аудіоданих. Асоціація метаданих стану обробки даних з аудіоданими є синхронною за часом. Таким чином, ці (останні прийняті або оновлені) метадані стану обробки даних вказують, що відповідні аудіодані одночасно містять результати обробки аудіоданих зазначеного типу (типів). У деяких випадках метадані стану обробки даних можуть містити історію обробки даних і/або деякі або всі параметри, які були використані при обробці даних зазначених типів і/або отримані при такій обробці даних. На додаток, метадані стану обробки даних можуть містити щонайменше одну ознаку, або характеристику, що відповідає аудіоданим, яка була обчислена або витягнута із цих аудіоданих. Метадані стану обробки даних можуть також містити інші метадані, що не відносяться або не отримані в результаті якої-небудь обробки відповідних аудіоданих. Наприклад, приватним модулем обробки аудіоданих для передачі іншим модулям обробки аудіоданих можуть бути додані дані третьої сторони, дані супроводу, ідентифікатори, відомості про власників або стандарти, дані користувацьких

55 коментарів, дані користувацьких переваг і т.д.

Усюди в даному розкритті, включаючи формулу винаходу, вираз «метадані стану обробки гучності» (або «LPSM») означає метадані стану обробки даних, що служать ознакою стану обробки гучності відповідних аудіоданих (наприклад, ознакою того, обробка гучності якого типу (типів) була виконана на цих аудіоданих), а також, як правило - щонайменше, однієї ознаки, або характеристики (наприклад, гучності), що відповідають аудіоданим. Метадані стану обробки

60

гучності можуть містити дані, що не є (тоді, коли вони розглядаються самі по собі) метаданими стану обробки гучності.

Усюди в даному розкритті, включаючи формулу винаходу, вираз «канал» (або «аудіоканал») означає монофонічний звуковий сигнал.

5 Усюди в даному розкритті, включаючи формулу винаходу, вираз «звукова програма» означає набір з одного або декількох аудіоканалів, а також, необов'язково, асоційовані метадані (наприклад, метадані, що описують необхідне просторове представлення звуку й/або PIM, i/або SSM, i/або LPSM, i/або метадані границь програми).

10 Усюди в даному розкритті, включаючи формулу винаходу, вираз «метадані границь програми» означає метадані кодованого бітового аудіопотоку, де цей кодований бітовий аудіопотік служить ознакою, щонайменше, однієї звукової програми (наприклад, двох або більшої кількості звукових програм), а граничні метадані програми служать ознакою місця розташування в бітовому потоці, щонайменше, однієї границі (початку й/або кінця), щонайменше, однієї зазначеної звукової програми. Наприклад, метадані границь програми (з 15 кодованого бітового аудіопотоку, що служить ознакою звукової програми) можуть містити метадані, що служать ознакою місця розташування (наприклад, початку «N»-го кадра бітового потоку або місця розташування «M»-го дискретного значення в «N»-м кадрі бітового потоку) початку цієї програми, а додаткові метадані служать ознакою місця розташування (наприклад, початку «J»-го кадра бітового потоку або місця розташування «K»-го дискретного значення в 20 «J»-м кадрі бітового потоку) кінця програми.

Усюди в даному розкритті, включаючи формулу винаходу, термін «зв'язується», або «зв'язаний», використовується, як такий, що позначає, або пряме, або непряме з'єднання. Так, якщо перший пристрій зв'язується із другим пристроєм, то з'єднання може здійснюватися через пряме з'єднання або через непряме з'єднання через інші пристрої й з'єднання.

25 Докладний опис варіантів здійснення винаходу

Типовий потік аудіоданих містить, як звуковий вміст (наприклад, один або кілька каналів звукового вмісту), так і метадані, що служать ознакою, щонайменше, однієї характеристики звукового вмісту. Наприклад, у бітовому потоці AC-3 є кілька параметрів метаданих аудіоданих, спеціально призначених для використання при зміні звучання програми, доставленої в 30 середовище для прослуховування. Одним із цих параметрів метаданих є параметр DIALNORM, призначений для вказівки середнього рівня діалогу у звуковій програмі й використовуваний для визначення рівня сигналу відтворення звуку.

У ході відтворення бітового потоку, що містить послідовність різних сегментів звукової програми (кожний з яких містить відмінний параметр DIALNORM), кодер AC-3 використовує 35 параметр DIALNORM кожного із сегментів для виконання обробки гучності того типу, який модифікує рівень відтворення, або гучність, так, щоб сприймана гучність діалогу із зазначеної послідовності сегментів перебувала на узгодженому рівні. Кожний кодований сегмент (елемент) аудіоданих у послідовності кодованих елементів аудіоданих міг би (взагалі) містити відрізний параметр DIALNORM, і декодер масштабував би рівень кожного із цих елементів так, щоб 40 рівень відтворення, або гучність, діалогу для кожного такого елемента був однаковий або дуже схожий, хоча це може зажадати застосування різних величин посилення до різних елементів у ході відтворення.

Як правило, DIALNORM встановлюється користувачем, і він не генерується автоматично, хоча існує обирне за замовчуванням значення DIALNORM якщо користувач не встановлює 45 ніяке значення. Наприклад, створювач вмісту може виконувати вимірювання гучності за допомогою пристрою, зовнішнього стосовно кодера AC-3, а потім передати результат (що виступає ознакою гучності мовного діалогу зі звукової програми) у кодер для встановлення значення DIALNORM. Таким чином, вірне встановлення параметра DIALNORM довіряється створювачу вмісту.

50 Є кілька різних причин того, чому параметр DIALNORM у бітовому потоці AC-3 може бути невірним. По-перше, кожний кодер AC-3 містить використовуване за замовчуванням значення DIALNORM, яке використовується в ході генерування бітового потоку, якщо створювач вмісту не встановив значення DIALNORM. Це використовуване за замовчуванням значення може суттєво відрізнитися від фактичного рівня гучності діалогу у звуковому сигналі. По-друге, навіть якщо 55 створювач вмісту вимірює гучність і відповідно встановлює значення DIALNORM, при цьому міг бути використаний алгоритм виміру гучності або вимірник, не відповідний до рекомендованого способу виміру гучності AC-3, що в результаті приводить до невірного значення DIALNORM. По-третє, навіть якщо бітовий потік AC-3 був створений зі значенням DIALNORM, вірно обмірюваним і встановленим створювачем вмісту, воно могло бути змінене на невірне значення 60 в ході передачі й/або зберігання цього бітового потоку. Наприклад, для телевізійних додатків

не є рідкістю декодування, модифікація, а потім повторне кодування бітових потоків AC-3 з використанням невірних відомостей про DIALNORM у метаданих. Таким чином, значення DIALNORM, укладене в бітовому потоці AC-3, може бути невірним або неточним і, таким чином, може впливати на якість вражень від прослуховування.

Крім того, параметр DIALNORM не вказує стан обробки гучності відповідних аудіоданих (наприклад, те, обробка гучності якого типу (типів) була виконана на цих аудіоданих). Метадані стану обробки гучності (у форматі, який передбачений у деяких варіантах здійснення справжнього винаходу) є корисними для полегшення адаптивної обробки гучності бітового аудіопотоку й/або перевірки вірогідності стану обробки гучності й гучності звукового вмісту особливо ефективним образом.

Незважаючи на те, що даний винахід не обмежений використанням з бітовим потоком AC-3, бітовим потоком E-AC-3, або бітовим потоком Dolby E, для зручності він буде описаний у варіантах здійснення, де він генерує, декодує або інакше обробляє такий бітовий потік.

Кодований бітовий потік AC-3 містить метадані й від одного до шести каналів звукового вмісту. Цей звуковий вміст являє собою аудіодані, які були стиснуті з використанням перцепційного звукового кодування. Зазначені метадані містять кілька параметрів метаданих аудіоданих, призначених для використання при зміні звучання програми, доставленої в середовище для прослуховування.

Кожний кадр кодованого бітового аудіопотоку AC-3 містить звуковий вміст і метадані для 1536 дискретних значень цифрових аудіоданих. При частоті дискретизації 48 кГц це являє собою 32 мілісекунди цифрового звуку або частоту 31,25 кадрів, що попадають на секунду звуку.

Кожний кадр кодованого бітового аудіопотоку E-AC-3 містить звуковий вміст і метадані для 256, 512, 768 або 1536 дискретних значень цифрових аудіоданих залежно від того, містить цей кадр, відповідно, один, два, три або шість блоків аудіоданих. При частоті дискретизації 48 кГц це становить, відповідно, 5,333, 10,667, 16 або 32 мілісекунд цифрового звуку або частоту, відповідно, 189,9, 93,75, 62,5 або 31,25 кадрів, що попадають на секунду звуку.

Як зазначено на Фіг. 4, кожен кадр AC-3 розділений на секції (сегменти), що містять: секцію відомостей про синхронізацію (SI), що містить (як показано на Фіг. 5) синхрослово (SW) і перше із двох слів виправлення помилок (CRC1); секцію відомостей про бітовий потік (BSI), що містить більшу частину метаданих; шість аудіоблоків (AB0-AB5), що містять стислі дані звукового вмісту (а також здатні містити метадані); сегменти зайвих бітів (W) (також відомі як «поля ігнорованих даних»), що містять які-небудь зайві біти, які залишилися після стиснення звукового вмісту; секцію допоміжних відомостей (AUX), яка також може містити метадані; і друге із двох слів виправлення помилок (CRC2).

Як зазначено на Фіг. 7, кожен кадр E-AC-3 розділений на секції (сегменти), що містять: секцію відомостей про синхронізацію (SI), що містить (як показано на Фіг. 5) синхрослово (SW); секцію відомостей про бітовий потік (BSI), що містить більшу частину метаданих; від одного до шести аудіоблоків (AB0-AB5), що містять стислі дані звукового вмісту (а також здатні включати метадані); сегменти зайвих бітів (W) (також відомі як «поля ігнорованих даних»), що містять які-небудь зайві біти, що залишилися після стиснення звукового вмісту (незважаючи на те, що показаний тільки один сегмент зайвих бітів, за кожним аудіоблоком, як правило, може слідувати відрізний сегмент зайвих бітів); секцію допоміжних відомостей (AUX), яка також може містити метадані; і слово виправлення помилок (CRC).

У бітовому потоці AC-3 (або E-AC-3) є кілька параметрів метаданих аудіоданих, спеціально призначених для використання при зміні звучання програми, доставленої в середовище для прослуховування. Одним з таких параметрів метаданих є параметр DIALNORM, ув'язнений у сегменті BSI.

Як показано на Фіг. 6, сегмент BSI кадра AC-3 містить п'ятибітний параметр («DIALNORM»), що вказує значення DIALNORM для цієї програми. П'ятибітний параметр («DIALNORM2»), що вказує значення DIALNORM для другої звукової програми, що переноситься в тому ж кадрові AC-3, включають, якщо режим звукового кодування («astmod») кадра AC-3 рівний «0», що вказує на те, що у вживанні перебуває подвійна монофонічна конфігурація каналів, або «1+1».

Сегмент BSI також містить прапор («addbsie»), що вказує на присутність (або відсутність) додаткових відомостей про бітовий потік, що слідує за бітом «addbsie», параметр («addbsil»), що вказує довжину яких-небудь додаткових відомостей про бітовий потік, що слідує за значенням «addbsil» і до 64 бітів додаткових відомостей про бітовий потік («addbsi»), що випливають за значенням «addbsil».

Сегмент BSI містить і інші значення метаданих, не показані конкретно на Фіг. 6.

Відповідно до одного із класів варіантів здійснення, кодований бітовий аудіопотік служить

ознакою декількох вкладених потоків звукового вмісту. У деяких випадках, ці вкладені потоки служать ознакою звукового вмісту багатоканальної програми, а кожний із вкладених потоків служить ознакою одного або декількох каналів цієї програми. В інших випадках, кілька вкладених потоків кодованого бітового аудіопотоку служать ознаками звукового вмісту декількох звукових програм, як правило, «головної» звукової програми (яка може являти собою багатоканальну програму) і, щонайменше, ще однієї звукової програми (наприклад, програми, що представляє собою коментарі до головної звукової програми).

Кодований бітовий аудіопотік, що служить ознакою, щонайменше, однієї звукової програми, неодмінно містить, щонайменше, один «незалежний» вкладений потік звукового вмісту. Цей незалежний вкладений потік служить ознакою, щонайменше, одного каналу звукової програми (наприклад, цей незалежний вкладений потік може служити ознакою п'яти каналів широкосмугових гучномовців традиційної 5.1-канальної звукової програми). У цьому документі ця звукова програма йменується «головною» програмою.

У деяких класах варіантів здійснення кодований бітовий аудіопотік служить ознакою двох або більшої кількості звукових програм («головної» програми й, щонайменше, ще однієї звукової програми). У таких випадках, цей бітовий потік містить два або більшу кількість незалежних вкладених потоків: перший незалежний вкладений потік, що служить ознакою, щонайменше, одного каналу головної програми; і, щонайменше, ще один незалежний вкладений потік, що служить ознакою, щонайменше, одного каналу іншої звукової програми (програми, окремої від головної програми). Кожний незалежний бітовий потік може бути декодований незалежно, і декодер може діяти для декодування тільки підмножини (а не всіх) незалежних вкладених потоків кодованого бітового потоку.

В одному з типових прикладів кодованого бітового аудіопотоку, що служить ознакою двох незалежних вкладених потоків, один із цих незалежних вкладених потоків служить ознакою каналів гучномовців стандартного формату багатоканальної звукової програми (наприклад, лівого, правого, центрального, лівого оточуючого, правого оточуючого каналів широкосмугових гучномовців 5.1-канальної головної програми), а інший незалежний вкладений потік служить ознакою монофонічного звукового коментаря до головної програми (наприклад, коментаря режисера кінофільму, де головна програма являє собою звукову доріжку цього кінофільму). В іншому прикладі кодованого бітового аудіопотоку, що служить ознакою декількох незалежних вкладених потоків, один із цих незалежних вкладених потоків служить ознакою каналів гучномовців стандартного формату багатоканальної головної програми (наприклад, 5.1-канальної головної програми), що містить діалог на першій мові (наприклад, ознакою діалогу може служити один з каналів гучномовців головної програми), а кожний наступний незалежний вкладений потік служить ознакою монофонічного перекладу цього діалогу (на іншу мову).

Необов'язково, кодований бітовий аудіопотік, що служить ознакою головної програми (а також, необов'язково, щонайменше, ще однієї звукової програми) містить, щонайменше, один «залежний» вкладений потік звукового вмісту. Кожний залежний вкладений потік асоційований з одним незалежним вкладеним потоком бітового потоку й служить ознакою, щонайменше, одного додаткового каналу програми (наприклад, головної програми), вміст якого вказується цим асоційованим незалежним вкладеним потоком (тобто зазначений залежний вкладений потік служить ознакою, щонайменше, одного каналу програми, не зазначеного асоційованим незалежним вкладеним потоком, а цей асоційований незалежний вкладений потік служить ознакою, щонайменше, одного каналу програми).

В одному із прикладів кодованого бітового потоку, що містить незалежний вкладений потік (що служить ознакою, щонайменше, одного каналу головної програми), цей бітовий потік також містить залежний вкладений потік (асоційований із цим незалежним вкладеним потоком), що служить ознакою одного або декількох каналів гучномовців головної програми. Такі додаткові канали гучномовців є додатковими до каналу (каналів) головної програми, що вказується незалежним вкладеним потоком. Наприклад, якщо незалежний вкладений потік служить ознакою стандартного формату лівого, правого, центрального, лівого навколишнього, правого навколишнього каналів широкосмугових гучномовців 7.1-канальної головної програми, то залежний вкладений потік може служити ознакою двох інших каналів широкосмугових гучномовців цієї головної програми.

У відповідності зі стандартом E-AC-3, бітовий потік E-AC-3 повинен служити ознакою, щонайменше, одного незалежного вкладеного потоку (наприклад, єдиного бітового потоку AC-3) і може служити ознакою до восьми незалежних вкладених потоків. Кожний незалежний вкладений потік бітового потоку E-AC-3 може бути асоційований з кількістю до восьми залежних вкладених потоків.

Бітовий потік E-AC-3 містить метадані, що служать ознакою структури вкладених потоків

бітового потоку. Наприклад, поле «chanmap» у секції відомостей про бітовий потік (BSI) бітового потоку E-AC-3 визначає схему каналів для каналів програми, що вказуються залежним вкладеним потоком цього бітового потоку. Однак метадані, що служать ознакою структури вкладених потоків, звичайно укладені в бітовому потоці E-AC-3 у такому форматі, щоб до них
 5 було зручно одержувати доступ і використовувати їх (у ході декодування кодованого бітового потоку E-AC-3) тільки за допомогою декодера E-AC-3; а не у форматі для доступу й використання після декодування (наприклад, при використанні постпроцесора) або перед декодуванням (наприклад, при використанні процесора, виконаного з можливістю розпізнавання метаданих). Також існує ризик того, що декодер може невірно ідентифікувати вкладені потоки
 10 традиційного кодованого бітового потоку E-AC-3, використовуючи метадані, включені традиційним образом, і до справжнього винаходу не було відомо, яким чином включати метадані структури вкладених потоків у кодований бітовий потік (наприклад, у кодований бітовий потік E-AC-3) у такому форматі, щоб уможливити зручне й ефективне виявлення й виправлення помилок в ідентифікації вкладених потоків у ході декодування зазначеного
 15 бітового потоку.

Бітовий потік E-AC-3 також може містити метадані, що відносяться до звукового вмісту звукової програми. Наприклад, бітовий потік E-AC-3, що служить ознакою звукової програми, містить метадані, що служать ознакою мінімальної й максимальної частот, до яких для
 20 кодування вмісту програми була застосована обробка розтягування спектра (і кодування зі зв'язуванням каналів). Однак такі метадані звичайно включені в бітовий потік E-AC-3 у такому форматі, що одержувати до них доступ і використовувати їх (у ході декодування кодованого бітового потоку E-AC-3) зручно тільки за допомогою декодера E-AC-3; а не за допомогою доступу й використання після декодування (наприклад, під час використання постпроцесора), або перед декодуванням (наприклад, під час використання процесора, виконаного з можливістю
 25 розпізнавання метаданих). Крім того, такі метадані не включені в бітовий потік E-AC-3 у форматі, який дозволяв би зручне й ефективне виявлення помилок і виправлення помилок ідентифікації таких метаданих у ході декодування бітового потоку.

Відповідно до типових варіантів здійснення винаходу, PIM і/або SSM (а також, необов'язково, інші метадані, наприклад, метадані стану обробки гучності, або «LPSM») вбудовують в одне або декілька зарезервованих полів (або областей) сегментів метаданих
 30 бітового аудіопотоку, що також містить аудіодані в інших сегментах (сегментах аудіоданих). Як правило, щонайменше один сегмент кожного кадра цього бітового потоку містить PIM або SSM, і, щонайменше, ще один сегмент цього кадра містить відповідні аудіодані (тобто аудіодані, структура вкладених потоків яких вказується за допомогою SSM, і/або аудіоданих, що мають, щонайменше, одну характеристику або властивість, що вказується PIM).

В одному із класів варіантів здійснення кожний сегмент метаданих являє собою структуру даних (іноді іменовану в цьому документі контейнером), здатну містити одну або кілька
 40 корисних навантажень метаданих. Кожне корисне навантаження включає заголовок, що містить індивідуальний ідентифікатор корисного навантаження (і конфігураційні дані корисного навантаження), що передбачає точно виражений показник типу метаданих, які присутні в цьому корисному навантаженні. Порядок корисних навантажень у контейнері є невизначеним, тому корисні навантаження можуть зберігатися в будь-якому порядку, і синтаксичний аналізатор повинен мати можливість виконувати синтаксичний аналіз усього контейнера для добування
 45 значимих корисних навантажень і зневаги тими з них, які не є значимими або є непідтримуваними. Фігура 8 (описувана нижче) ілюструє структуру такого контейнера й корисних навантажень у контейнері.

Повідомлення метаданих (наприклад, SSM і/або PIM, і/або LPSM) за ланцюгом обробки аудіоданих є особливо корисним тоді, коли двом або більшій кількості модулів обробки аудіоданих необхідно працювати спільно один з одним усюди в ланцюгу обробки даних (або
 50 протягом усього життєвого циклу вмісту). Під час відсутності включення метаданих у бітовий аудіопотік можуть виникати серйозні труднощі обробки мультимедійних даних, такі, як погіршення якості, рівня або просторові погіршення, наприклад, тоді, коли два або більшу кількість аудіокодеків використовують у ланцюгу й однобічне регулювання рівня гучності застосовується на шляху бітового потоку до пристрою споживання мультимедійних даних (або
 55 до точки представлення звукового вмісту бітового потоку) більш одного разу.

Метадані стану обробки гучності (LPSM), впроваджені в бітовий аудіопотік відповідно до деяких варіантів здійснення винаходу, можуть бути аутентифіковані й перевірені на вірогідність, наприклад, щоб дозволити регулятивним органам гучності, перевіряти, чи перебуває гучність конкретної програми вже в заданих межах, і що відповідні аудіодані самі по собі не були
 60 модифіковані (за допомогою чого забезпечується відповідність застосовним нормам). Для цієї

перевірки, замість обчислення гучності заново, може бути зчитане значення гучності, укладене в блоці даних, що містить метадані стану обробки гучності. У відповідь на LPSM регулятивний орган може визначати, що відповідний звуковий вміст перебуває у відповідності (що зазначене LPSM) із законодавчими й/або нормативними вимогами (наприклад, з нормами, оприлюдненими в Законі про зменшення гучності комерційних рекламних оголошень (Commercial Advertisement Loudness Mitigation Act, також відомому як закон «CALM»)) без необхідності в обчисленні гучності звукового вмісту.

Фіг. 1 являє собою блок-схему одного із прикладів ланцюга обробки аудіоданих (системи обробки аудіоданих), у якій один або декілька з елементів системи можуть бути виконані відповідно до одного з варіантів здійснення цього винаходу. Ця система містить наступні елементи, зв'язані один з одним так, як це показано: модуль попередньої обробки даних, кодер, модуль аналізу сигналу й виправлення метаданих, перетворювач коду, декодер і модуль попередньої обробки даних. У варіантах показаної системи пропущені один або декілька із цих елементів або додані додаткові модулі обробки аудіоданих.

У деяких реалізаціях модуль попередньої обробки даних за Фіг. 1 виконаний з можливістю приймання в якості введення дискретних значень PCM (у часовій області), що містять звуковий вміст, і виводу оброблених дискретних значень PCM. Кодер може бути виконаний з можливістю приймання у якості введення дискретних значень PCM і виводу кодованого (наприклад, стислого) бітового аудіопотоку, що служить ознакою звукового вмісту. Дані бітового потоку, що служать ознакою звукового вмісту, іноді йменуються в цьому документі «аудіоданими». Якщо кодер виконаний відповідно до типового варіанта здійснення цього винаходу, то вивід бітового потоку з кодера крім аудіоданих містить PIM і/або SSM (а також, необов'язково метадані стану обробки гучності й/або інші метадані).

Модуль аналізу сигналу й виправлення метаданих за Фіг. 1 може приймати в якості вводу один або декілька кодованих бітових аудіопотоків і визначати (наприклад, перевіряти на вірогідність), чи є вірними метадані (наприклад, метадані стану обробки даних) у кожному кодованому бітовому потоці, шляхом виконання аналізу сигналу (наприклад, використовуючи метадані границь програми в кодованому бітовому аудіопотоці). Якщо модуль аналізу сигналу й виправлення метаданих визначає, що включені метадані не є достовірними, він, як правило, заміщає невірне значення вірним значенням (значеннями), отриманим(и) з аналізу сигналу. Таким чином, вивід кожного кодованого бітового аудіопотоку з модуля аналізу сигналу й виправлення метаданих може містити виправлені (або не виправлені) метадані стану обробки даних, а також кодовані аудіодані.

Перетворювач коду за Фіг. 1 може ухвалювати в якості введення кодовані бітові аудіопотоки й виводити у відповідь модифіковані (наприклад, кодовані іншим образом) бітові аудіопотоки (наприклад, шляхом декодування вхідного потоку й повторного кодування цього декодованого потоку в іншому форматі кодування). Якщо перетворювач коду виконаний відповідно до одного з типових варіантів здійснення справжнього винаходу, то зазначений вивід бітового аудіопотоку з перетворювача коду поряд з кодованими аудіоданими містить SSM і/або PIM (а також, як правило, інші метадані). Ці метадані могли бути укладені у вхідному бітовому потоці.

Декодер за Фіг. 1 може приймати в якості введення кодовані (наприклад, стислі) бітові аудіопотоки й (у відповідь) виводити потоки декодованих дискретних значень PCM аудіоданих. Якщо декодер виконаний відповідно до одного з типових варіантів здійснення цього винаходу, вивід цього декодера при типовій роботі являє собою або містить що-небудь із наступного:

- потік дискретних значень аудіоданих і, щонайменше, один відповідний потік SSM і/або PIM (а також, як правило, інших метаданих), витягнуті із вхідного кодованого бітового потоку; або
- потік дискретних значень аудіоданих і відповідний потік керуючих бітів, визначених на основі SSM і/або PIM (а також, як правило, інших метаданих, наприклад, LPSM), витягнутих із вхідного кодованого бітового потоку; або

- потік дискретних значень аудіоданих без відповідного потоку метаданих або керуючих бітів, визначених на основі метаданих. У цьому, останньому випадку, декодер може витягати метадані із вхідного кодованого бітового потоку й виконувати на цих витягнутих метаданих, щонайменше, одну операцію (наприклад, перевірку вірогідності), навіть якщо він не виводить витягнуті метадані або визначені на основі них керуючі біти.

При виконанні модуля заключної обробки даних за Фіг. 1 відповідно до одного з типових варіантів здійснення цього винаходу, цей модуль заключної обробки даних є виконанням з можливістю приймання потоку декодованих дискретних значень PCM аудіоданих і виконання на них заключної обробки даних (наприклад, регулювання рівня гучності звукового вмісту) з використанням SSM і/або PIM (а також, як правило, інших метаданих, наприклад, LPSM), прийнятих разом із цими дискретними значеннями, або керуючих бітів, визначених декодером

на основі метаданих, прийнятих разом із цими дискретними значеннями. Модуль заключної обробки даних, як правило, також виконаний з можливістю представлення звукового вмісту для відтворення одним або декількома гучномовцями, що підлягає заключній обробці.

Типові варіанти здійснення цього винаходу передбачають удосконалений ланцюг обробки аудіоданих, у якому модулі обробки аудіоданих (наприклад, кодери, декодери, перетворювачі коду, а також модулі попередньої обробки й заключної обробки даних) пристосовують відповідну їм обробку даних для застосування на аудіоданих відповідно до одночасного стану медіаданих, яке, відповідно, вказується метаданими, прийнятими зазначеними модулями обробки аудіоданих.

Уведення аудіоданих у який-небудь модуль обробки аудіоданих системи за Фіг. 1 (наприклад, у кодер або перетворювач коду за Фіг. 1) може містити поряд з аудіоданими (наприклад, з кодованими аудіоданими) SSM і/або PIM (а також, необов'язково, інші метадані). Ці метадані могли бути включені в введення аудіоданих іншим елементом системи за Фіг. 1 (або іншим джерелом, не показаним на Фіг. 1) відповідно до одного з варіантів здійснення справжнього винаходу. Модуль обробки даних, що ухвалює введення аудіоданих (з метаданими) може бути виконаний з можливістю виконання, щонайменше, однієї операції (наприклад, перевірки вірогідності) на цих метаданих або у відповідь на ці метадані (наприклад, адаптивної обробки вхідних аудіоданих), а також, як правило, містить у своїх вихідних аудіоданих метадані, оброблену версію цих метаданих або керуючі біти, визначені на основі цих метаданих.

Один з типових варіантів здійснення винахідницького модуля обробки аудіоданих (або процесора аудіоданих) виконаний з можливістю виконання адаптивної обробки аудіоданих на основі стану аудіоданих, що вказується метаданими, що відповідають цим аудіоданим. У деяких варіантах здійснення адаптивна обробка даних являє собою (або включає) обробку гучності, якщо метадані вказують, що обробка гучності або подібна до неї обробка даних не була вже виконана на цих аудіоданих, але не являє собою (і не включає) обробку гучності, якщо метадані вказують, що така обробка гучності або подібна до неї обробка даних уже була виконана на цих аудіоданих. У деяких варіантах здійснення адаптивна обробка даних являє собою або включає перевірку вірогідності метаданих (наприклад, виконувати в підмодулі перевірку вірогідності метаданих), забезпечуючи те, що модуль обробки аудіоданих виконує іншу адаптивну обробку на аудіоданих на основі стану аудіоданих, що вказується метаданими. У деяких варіантах здійснення перевірка вірогідності визначає надійність метаданих, асоційованих з аудіоданими (або таких, що містяться спільно з ними у бітовому потоці). Наприклад, якщо перевірені на вірогідність метадані є надійними, то результати раніше виконаної обробки аудіоданих одного з типів можна використовувати повторно, уникаючи нового виконання обробки аудіоданих того ж типу. З іншого боку, якщо виявлено, що метадані було підроблено (або вони є ненадійними в інших відносинах), у такому разі обробку метаданих цього типу, ймовірно, виконано раніше (що вказується ненадійними метаданими), можна повторити за допомогою модуля обробки аудіоданих, і/або модулем обробки аудіоданих на метаданих і/або аудіоданих може бути виконана інша обробка даних. Модуль обробки аудіоданих також може бути виконаний з можливістю сигналізації іншим модулям обробки аудіоданих у спадному напрямку вдосконаленого ланцюга обробки медіаданих про те, що метадані (наприклад, присутні в бітовому потоці медіаданих) є достовірними (наприклад, заснованими на збігу витягнутої криптографічної величини й контрольної криптографічної величини).

Фіг. 2 являє собою блок-схему кодера (100), що являє собою один з варіантів здійснення винахідницького модуля обробки аудіоданих. Кожний з компонентів або елементів кодера 100 може бути реалізований як один або кілька процесів і/або одна або кілька схем (наприклад, мікросхем ASIC, матриць FPGA або інших інтегральних мікросхем), в апаратному забезпеченні, програмному забезпеченні або в комбінації апаратного й програмного забезпечення. Кодер 100 містить буфер 110 кадрів, синтаксичний аналізатор 111, декодер 101, засіб 102 перевірки вірогідності стану аудіоданих, щабель 103 обробки гучності, щабель 104 вибору аудіопотоку, кодер 105, щабель 107 формування швидкості передачі даних/засобу форматування, щабель 106 генерування метаданих, підсистему 108 виміру гучності діалогу й буфер 109 кадрів, з'єднані так, як це показано. Як правило, кодер 100 також містить інші елементи обробки даних (не показані).

Кодер 100 (що представляє собою перетворювач коду) виконаний з можливістю перетворення вхідного бітового аудіопотоку (який, наприклад, може являти собою бітовий потік AC-3, бітовий потік E-AC-3 або бітовий потік Dolby E) у кодований вихідний бітовий потік (який, наприклад, може являти собою інший бітовий потік, обраний з бітового потоку AC-3, бітового потоку E-AC-3 або бітового потоку Dolby E), що полягає у виконанні адаптивної й

автоматизованої обробки гучності з використанням метаданих стану обробки гучності, які містяться у вхідному бітовому потоці. Наприклад, кодер 100 може бути виконаний з можливістю перетворення вхідного бітового потоку Dolby E (формат, як правило, використовуваний у виробничому й віщальному устаткуванні, але не в споживчих пристроях, які приймають звукові програми, що віщають на них) у кодований вихідний бітовий аудіопотік (придатний для віщання на споживчі пристрої) у форматі AC-3 або E-AC-3.

Система за Фіг. 2 також містить підсистему 150 доставки кодованих аудіоданих (яка зберігає в пам'яті й/або доставляє кодовані бітові потоки, що виходять із кодера 100) і декодер 152. Кодований бітовий аудіопотік, що виходить із кодера 100, може бути збережений у пам'яті підсистеми 150 (наприклад, у формі диска DVD або Blu-ray) або переданий підсистемою 150 (яка може реалізовувати канал або мережу зв'язку), або він може бути, як збережений, так і переданий підсистемою 150. Декодер 152 виконаний з можливістю декодування кодованого бітового аудіопотоку (що генерується кодером 100), який він приймає за допомогою підсистеми 150, що оснований на добуванні метаданих (PIM і/або SSM, а також, необов'язково, метаданих стану обробки гучності й/або інших метаданих) з кожного кадра бітового потоку (а також, необов'язково, на добуванні з бітового потоку метаданих границь програми) і в генеруванні декодованих аудіоданих. Як правило, декодер 152 виконаний з можливістю виконання адаптивної обробки даних на декодованих аудіоданих з використанням PIM і/або SSM, і/або LPSM (а також, необов'язково, метаданих границь програми), і/або спрямовування декодованих аудіоданих і метаданих у постпроцесор, виконаний з можливістю виконання адаптивної обробки даних на декодованих аудіоданих з використанням метаданих. Як правило, декодер 152 містить буфер, який зберігає в пам'яті (наприклад, неминущим чином) кодований аудіопотік, прийнятий з підсистеми 150.

Різні реалізації кодера 100 і декодера 152 виконані з можливістю виконання різних варіантів здійснення способу за винаходом.

Буфер 110 кадрів являє собою буферну пам'ять, пов'язану із прийманням кодованого вхідного бітового аудіопотоку. У дії, буфер 110 зберігає (наприклад, неминущим чином), щонайменше, один кадр кодованого бітового аудіопотоку, а послідовність кадрів кодованого бітового аудіопотоку направляється з буфера 110 у синтаксичний аналізатор 111.

Синтаксичний аналізатор 111 зв'язаний і виконаний з можливістю добування PIM і/або SSM і метаданих стану обробки гучності (LPSM), а також, необов'язково, метаданих границь програми (і/або інших метаданих) з кожного кадра кодованих вхідних аудіоданих, у якому укладені ці метадані, для спрямовування, щонайменше, LPSM (а також, необов'язково, метаданих границь програми й/або інших метаданих) у засіб 102 перевірки вірогідності стану метаданих, на щабель 103 обробки гучності, щабель 106 і в підсистему 108 для добування аудіоданих з кодованих вхідних аудіоданих і для напрямку аудіоданих у декодер 101. Декодер 101 кодера 100 виконаний з можливістю декодування цих аудіоданих з метою генерування декодованих аудіоданих і спрямовування цих декодованих аудіоданих на щабель 103 обробки гучності, щабель 104 вибору аудіопотоку, у підсистему 108, а також, як правило, у засіб 102 перевірки вірогідності стану.

Засіб 102 перевірки вірогідності стану виконаний з можливістю аутентифікації й перевірки вірогідності метаданих LPSM, що направляються в нього (і, необов'язково, інших метаданих). У деяких варіантах здійснення метадані LPSM являють собою блок даних (або укладені в нього), який був укладений у вхідному бітовому потоці (наприклад, відповідно до одного з варіантів здійснення справжнього винаходу). Цей блок може містити значення криптографічної хеш-функції (хеш-коду аутентифікації повідомлень, або «HMAC») для обробки LPSM (а також, необов'язково, інших метаданих) і/або аудіоданих, що лежать у їхній основі (доставлених з декодера 101 у засіб 102 перевірки вірогідності). У цих варіантах здійснення блок даних може містити цифровий підпис, тому модуль обробки аудіоданих у спадному напрямку може відносно легко аутентифікувати і перевіряти вірогідність зазначених метаданих стану обробки даних.

Наприклад, HMAC використовують для генерування згортки, і захисна величина (величини), ув'язнена у бітовому потоці за винаходом, може містити цю згортку. Зазначену згортку для кадра AC-3 можна генерувати в наступний спосіб.

1. Після кодування даних AC-3 і LPSM, байти даних кадра (зчеплені frame_data #1 і frame_data #2) і байти даних LPSM використовують у якості вводу для хеш-функції HMAC. Інші дані, які можуть бути присутніми у полі auxdata, при обчисленні згортки не враховують. Ці інші дані можуть являти собою байти, що не належать ані до даних AC-3, ані до даних LPSM. При обчисленні згортки HMAC можуть не враховуватися біти захисту, укладені в LPSM.

2. Після обчислення згортки, її записують у бітовий потік у поле, зарезервоване для бітів захисту.

3. Останнім етапом генерування повного кадра AC-3 є обчислення критерію CRC. Він записується в самий кінець кадра, і в розрахунки приймаються всі дані, що належать цьому кадрові, у тому числі біти LPSM.

Для перевірки вірогідності LPSM і/або інших метаданих (наприклад, у засобі 102 перевірки вірогідності) з метою забезпечення захищеної передачі й приймання метаданих і/або лежачих у їхній основі аудіоданих, можна використовувати й інші криптографічні методи, у тому числі, без обмеження, кожний з одного або декількох криптографічних методів, що не відносяться до HMAC. Наприклад, перевірку вірогідності (що використовує такий криптографічний метод) можна виконувати в кожному модулі обробки аудіоданих, що приймає один з варіантів здійснення винахідницького бітового аудіопотоку, для визначення того, чи були метадавні й відповідні аудіодані, вкладені в цей бітовий потік, піддані спеціальній обробці даних (або чи є вони її результатом), що вказується метаданими, і чи були вони модифіковані після виконання зазначеної спеціальної обробки даних.

Засіб 102 перевірки вірогідності стану направляє керуючі дані на щабель 104 вибору аудіопотоку, у генератор 106 метаданих і в підсистему 108 виміру гучності діалогу з метою вказання результатів операції перевірки вірогідності. У відповідь на ці керуючі дані щабель 104 може вибирати (і пропускати в кодер 105) одне з наступного:

- адаптивно оброблений вивід щабля 103 обробки гучності (наприклад, тоді, коли метадавні LPSM указують, що вивід аудіоданих з декодера 101 не був підданий обробці гучності спеціального типу, а керуючі біти із засобу 102 перевірки вірогідності вказують, що метадавні LPSM є достовірними); або

- вивід аудіоданих з декодера 101 (наприклад, тоді, коли метадавні LPSM указують, що вивід аудіоданих з декодера 101 уже був підданий обробці гучності спеціального типу, яка могла бути виконана щаблем 103, а керуючі біти із засобу 102 перевірки вірогідності вказують, що метадавні LPSM є достовірними).

Щабель 103 кодера 100 виконаний з можливістю виконання адаптивної обробки гучності на виводі декодованих аудіоданих з декодера 101 на основі однієї або декількох характеристик аудіоданих, що вказуються метаданими LPSM, витягнутими декодером 101. Щабель 103 може являти собою процесор керування гучністю й динамічним діапазоном у реальному часі в області перетворення. Щабель 103 може ухвалювати користувацьке введення (наприклад, цільові користувацькі значення гучності/динамічного діапазону або значення dialnorm) або введення інших метаданих (наприклад, даних третьої сторони одного або декількох типів, дані супроводу, ідентифікатори або відомості про стандарт, дані користувацьких коментарів, дані користувацьких переваг і т.д.) і/або інше введення (наприклад, із процесу обчислення контрольних сум) і використовувати це введення для обробки виводу декодованих аудіоданих з декодера 101. Щабель 103 може виконувати адаптивну обробку гучності на декодованих аудіоданих (вихідних з декодера 101), що служать ознакою єдиної звукової програми (що вказується метаданими границь програми, витягнутими синтаксичним аналізатором 111) і може скидати обробку гучності у відповідь на приймання декодованих аудіоданих (вихідних з декодера 101), що служать ознакою іншої звукової програми, що вказується метаданими границь програми, витягнутими синтаксичним аналізатором 111.

Підсистема 108 виміру гучності діалогу може діяти для визначення гучності сегментів декодованих аудіоданих (з декодера 101), що служать ознаками діалогу (або іншої мови), наприклад, використовуючи метадавні LPSM (і/або інші метадавні), витягнуті декодером 101, коли управляючі біти із засобу 102 перевірки вірогідності вказують, що метадавні LPSM є недостовірними. Дію підсистеми 108 виміру гучності діалогу можна скасувати, якщо метадавні LPSM указують раніше визначену гучність діалогу (або іншої мови) сегментів декодованих аудіоданих (з декодера 101), коли управляючі біти із засобу 102 перевірки вірогідності вказують, що метадавні LPSM є достовірними. Підсистема 108 може виконувати вимір гучності на декодованих аудіоданих, що служать ознакою єдиної звукової програми (що вказується метаданими границь програми, витягнутими синтаксичним аналізатором 111) і може скидати цей вимір у відповідь на приймання декодованих аудіоданих, що служать ознакою іншої звукової програми, що вказується метаданими границь програми.

Для зручного й простого вимірювання рівня діалогу у звуковому вмісті існують корисні інструментальні засоби (наприклад, вимірник гучності Dolby LM100). Деякі варіанти здійснення модуля APU за винаходом (наприклад, щаблі 108 кодера 100) реалізовані так, щоб вони містили такий інструментальний засіб (або виконували його функції) для виміру середньої гучності діалогу звукового вмісту з бітового аудіопотоку (наприклад, декодованого бітового потоку AC-3, що направляється на щабель 108 з декодера 101 кодера 100).

Якщо щабель 108 реалізований для виміру дійсної середньої гучності діалогу аудіоданих,

цей вимір може включати етап відділення сегментів звукового вмісту, що переважно містять мову. Сегменти аудіоданих, що переважно представляють собою мову, потім обробляються відповідно до алгоритму виміру гучності. Для аудіоданих, декодованих з бітового потоку AC-3, цей алгоритм може являти собою стандартний захід гучності, зважений по кривій K (відповідно до міжнародного стандарту ITU-R BS.1770).

У якості альтернативи, можна використовувати й інші заходи гучності (наприклад, заходи, засновані на психоакустичних моделях гучності).

Відділення мовних сегментів не є істотним для виміру середньої гучності діалогу аудіоданих. Однак воно підвищує точність заходу й, як правило, забезпечує більш задовільні результати з погляду слухача. Через те, що не весь звуковий уміст містить діалог (мова), захід гучності всього звукового вмісту може забезпечувати достатнє наближення рівня діалогу аудіоданих, що містили присутню в них мову.

Генератор 106 метаданих генерує (і/або пропускає на щабель 107) метадані, що підлягають включенню щаблем 107 у кодований потік, що підлягає виводу з кодера 100. Генератор 106 метаданих може пропускати на щабель 107 метадані LPSM (а також, необов'язково, LIM і/або PIM, і/або метадані границь програми, і/або інші метадані), витягнуті декодером 101 і/або синтаксичним аналізатором 111 (наприклад, коли управляючі біти із засобу 102 перевірки вірогідності вказують, що LPSM і/або інші метадані є достовірними), або генерувати нові метадані LIM і/або PIM, і/або LPSM, і/або інші метадані й направляти ці нові метадані на щабель 107 (наприклад, коли управляючі біти із засобу 102 перевірки вірогідності вказують, що метадані, витягнуті декодером 101, є недостовірними), або він може направляти на щабель 107 комбінацію метаданих, витягнутих декодером 101 і/або синтаксичним аналізатором 111, і заново згенерованих метаданих. Генератор 106 метаданих може включати дані гучності, згенеровані підсистемою 108, і, щонайменше, одну величину, що служить ознакою типу обробки гучності, виконаної підсистемою 108, у метадані LPSM, які він направляє на щабель 107 для включення в кодований бітовий потік, що підлягає виводу з кодера 100.

Генератор 106 метаданих може генерувати біти захисту (які можуть складатися з хеш-коду аутентифікації повідомлень, або «HMAC», або містити цей код), придатні для, щонайменше, однієї з наступних дій: розшифрування, аутентифікації або перевірки вірогідності метаданих LPSM (а також, необов'язково, інших метаданих), що підлягають включенню в кодований бітовий потік, і/або аудіоданих, що лежать у їхній основі, що підлягають включенню в цей кодований бітовий потік. Генератор 106 метаданих може доставляти ці біти захисту на щабель 107 для включення в кодований бітовий потік.

При типовій роботі підсистема 108 виміру гучності діалогу обробляє вивід аудіоданих з декодера 101 з метою генерування у відповідь на них значень гучності (наприклад, стробованих або нестробованих значень гучності діалогу) і значень динамічного діапазону. У відповідь на ці значення генератор 106 метаданих генерує метадані стану обробки гучності (LPSM) для включення (формування швидкості передачі даних/засобом форматування 107) у кодований бітовий потік, що підлягає виводу з кодера 100.

На додаток, необов'язково або в якості альтернативи, підсистеми 106 і/або 108 кодера 100 можуть виконувати додатковий аналіз аудіоданих для генерування метаданих, що служать ознаками, щонайменше, однієї характеристики аудіоданих, для включення в кодований бітовий потік, що підлягає виводу із щабля 107.

Кодер 105 кодує (наприклад, виконуючи на ньому стискання) вивід аудіоданих із щабля 104 вибору й направляє ці кодовані аудіодані на щабель 107 для включення в кодований бітовий потік, що підлягає виводу із щабля 107.

Щабель 107 ущільнює кодовані аудіодані з кодера 105 і метадані (утримуючі PIM і/або SSM) з генератора 106 для генерування кодованого бітового потоку, що підлягає виводу із щабля 107, переважно так, щоб цей кодований бітовий потік мав формат, визначений одним із кращих варіантів здійснення справжнього винаходу.

Буфер 109 кадрів являє собою буферну пам'ять, яка зберігає (наприклад, неминущим образом), щонайменше, один кадр із виводу кодованого бітового аудіопотоку із щабля 107, і послідовність кадрів кодованого бітового аудіопотоку потім направляється з буфера 109 як вивід кодера 100 у систему 150 доставки.

Метадані LPSM, що генеруються генератором 106 метаданих і включені в кодовані бітовий потік щаблем 107, як правило, служать ознаками стану обробки гучності відповідних аудіоданих (наприклад, того, обробка гучності якого типу (типів) була виконана на цих аудіоданих) і гучності (наприклад, обмірюваної гучності діалогу, стробованої і/або нестробованої гучності, і/або динамічного діапазону) відповідних аудіоданих.

У цьому документі «стробування» вимірів гучності й/або рівня, виконуване на аудіоданих,

відноситься до спеціального граничного значення рівня, або гучності, коли обчислене значення (значення), що перевищують це граничне значення, включаються в остаточний вимір (наприклад, шляхом пропуску короткострокових значень гучності нижче -60 дБ повної шкали в остаточних обмірюваних значеннях). Стробування на абсолютному значенні відноситься до

5 фіксованого рівня, або гучності, у той час як стробування на відносному значенні відноситься до значення, що не залежить від поточного «нестробованого» значення виміру.

У деяких реалізаціях кодера 100 кодований бітовий потік, буферований у пам'яті 109 (і виведений у систему 150 доставки), являє собою бітовий потік AC-3 або бітовий потік E-AC-3 і містить сегменти аудіоданих (наприклад, сегменти AB0-AB5 кадру, показаного на фіг. 4) і

10 сегменти метаданих, де сегменти аудіоданих служать ознаками аудіоданих, а кожний з щонайменше деяких із сегментів метаданих містить PIM і/або SSM (а також, необов'язково, інші метадані). Щабель 107 вставляє сегменти метаданих (що містять метадані) у бітовий потік у наступному форматі. Кожний із сегментів метаданих, що містять PIM і/або SSM, включається в

15 4 або фіг. 7) або в поле «addbsi» сегмента відомостей про бітовий потік («BSI»), або в поле auxdata (наприклад, у сегмент AUX, показаний на фіг. 4 або фіг. 7) наприкінці кадру цього бітового потоку. Кадр бітового потоку може містити один або два сегменти метаданих, кожний з яких містить метадані, і якщо цей кадр містить два сегменти метаданих, то один може бути у полі addbsi кадру, а інший - у полі AUX кадру.

20 У деяких варіантах здійснення кожний сегмент метаданих (іноді іменований у цьому документі «контейнером»), що вставляється шаблоном 107, має формат, що включає заголовок сегмента метаданих (а також, необов'язково, інші обов'язкові, або «базові», елементи) і одне або кілька корисних завантажень метаданих, що слідує за заголовком сегмента метаданих. Метадані SIM, якщо вони присутні, включаються в одне із корисних завантажень метаданих (що

25 ідентифікується за допомогою заголовка корисного завантаження і, як правило, що має формат першого типу). Метадані PIM, якщо вони присутні, включаються в ще одне з корисних завантажень метаданих (що ідентифікується за допомогою заголовка корисного завантаження і, як правило, що має формат другого типу). Аналогічно, метадані будь-якого іншого типу (якщо вони присутні) включаються в ще одне з корисних завантажень метаданих (що ідентифікується

30 за допомогою заголовка корисного завантаження і, як правило, що має формат, специфічний для цього типу метаданих). Цей ілюстративний формат уможливорює зручний доступ до SSM, PIM та інших метаданих в інші моменти часу, аніж під час декодування (наприклад, при використанні постпроцесора слідом за декодуванням або при використанні процесора, виконаного з можливістю розпізнавання метаданих без виконання повного декодування на

35 кодованому бітовому потоці), і уможливорює зручне й ефективне виявлення й виправлення помилок (наприклад, ідентифікації вкладених потоків) у ході декодування бітового потоку. Наприклад, за відсутністю доступу до SSM в ілюстративному форматі декодер може невірно ідентифікувати вірну кількість вкладених потоків, асоційованих із програмою. Одне корисне завантаження метаданих у сегменті метаданих може містити SSM, інше корисне завантаження метаданих може містити PIM, а також, необов'язково щонайменше ще одне корисне завантаження метаданих у сегменті метаданих може містити інші метадані (наприклад, метадані стану обробки гучності, або «LPSM»).

У деяких варіантах здійснення корисне завантаження метаданих структури вкладених потоків (SSM), що включається (шаблом 107) у кадр кодованого бітового потоку (наприклад, бітового потоку E-AC-3, що служить ознакою щонайменше однієї звукової програми), містить SSM у наступному форматі:

- заголовок корисного завантаження, як правило, що містить щонайменше одну величину-ідентифікатор (наприклад, 2-бітну величину, що служить ознакою версії формату SSM, а також, необов'язково значень довжини, періоду, лічильника й асоціації вкладених потоків); і,

50 - після заголовка - метадані незалежних вкладених потоків, що служать ознакою кількості незалежних вкладених потоків програми, що вказується цим бітовим потоком; і

- метадані залежних вкладених потоків, що служать ознакою того, чи містить кожний незалежний вкладений потік програми щонайменше один асоційований залежний вкладений потік (тобто, того, чи асоційований щонайменше один залежний вкладений потік із зазначеним кожним незалежним вкладеним потоком), і, якщо це так - кількості залежних вкладених потоків, асоційованих з кожним незалежним вкладеним потоком програми.

60 Передбачається, що незалежний вкладений потік кодованого бітового потоку може служити ознакою набору каналів гучномовців звукової програми (наприклад, каналів гучномовців звукової програми з 5.1 каналів гучномовців), і що кожний з одного або декількох вкладених потоків (асоційованих із зазначеним незалежним вкладеним потоком, що вказується

метаданими залежних вкладених потоків) може служити ознакою об'єктного каналу програми. Як правило, однак, незалежний вкладений потік кодованого бітового потоку служить ознакою набору каналів гучномовців програми, а кожний залежний вкладений потік, асоційований із цим незалежним вкладеним потоком (що вказується метаданими залежних вкладених потоків) служить ознакою щонайменше одного додаткового каналу програми.

У деяких варіантах здійснення корисне завантаження метаданих відомостей про програму (PIM), що включаються (щаблем 107) у кадр кодованого бітового потоку (наприклад, кодованого бітового потоку E-AC-3, що служить ознакою щонайменше однієї звукової програми) має наступний формат:

- заголовок корисного завантаження, як правило, що містить щонайменше одну величину-ідентифікатор (наприклад, величину, що служить ознакою версії формату PIM, а також, необов'язково значень довжини, періоду, лічильника й асоціації вкладених потоків); і,
- після заголовка - PIM у наступному форматі:

- метадані активних каналів, що служать ознакою кожного заглушеного каналу й кожного незаглушеного каналу звукової програми (тобто, того, який канал (канали) програми містить звукову інформацію, а який (якщо він присутній) містить тільки мовчання (як правило, протягом часу певної тривалості)). У варіантах здійснення, де кодований бітовий потік являє собою бітовий потік AC-3 або E-AC-3, метадані активних каналів у кадрі бітового потоку можна використовувати в комбінації з додатковими метаданими бітового потоку (наприклад, з полем режиму звукового кодування («astmod») кадру й, якщо воно присутнє, з полем chanmap у цьому кадрі або в кадрі (кадрах) асоційованих вкладених потоків)) для визначення того, який канал (канали) програми містить звукову інформацію, а який містить мовчання. Поле «astmod» кадру AC-3 або E-AC-3 вказує кількість ширококугових каналів звукової програми, що вказується звуковим вмістом кадру (наприклад, те, чи є ця програма 1.0-канальною монофонічною програмою, 2.0-канальною стереофонічною програмою або програмою, що містить ширококугові канали L, R, C, Ls, Rs), або те, що цей кадр служить ознакою двох незалежних 1.0-канальних монофонічних програм. Поле «chanmap» бітового потоку E-AC-3 вказує схему каналів для залежного вкладеного потоку, що вказується бітовим потоком. Метадані активних каналів можуть бути корисними для реалізації підвищувального мікшування (у постпроцесорі) у спадному напрямку відносно декодера, наприклад, для додавання звуку в канали, що містять мовчання на виводі декодера;

- метадані стану обробки знижувального мікшування, які служать ознакою того, чи зазнала ця програма знижувального мікшування (перед кодуванням або в ході його), і, якщо це так - того, знижувальне мікшування якого типу застосовувалося. Метадані стану обробки знижувального мікшування можуть бути корисними для реалізації підвищувального мікшування (у постпроцесорі) у спадному напрямку щодо декодера, наприклад, для підвищувального мікшування звукового вмісту програми з використанням параметрів, які найбільше відповідають типу застосованого знижувального мікшування. У тих варіантах здійснення, де кодований бітовий потік являє собою бітовий потік AC-3 або E-AC-3, метадані стану обробки знижувального мікшування можна використовувати в комбінації з полем режиму звукового кодування («astmod») кадру для визначення типу знижувального мікшування (якщо воно мало місце), застосованого до каналу (каналів) програми;

- метадані стану обробки підвищувального мікшування, які служать ознакою того, чи зазнала програма підвищувального мікшування (наприклад, від меншої кількості каналів) перед або під час кодування, і, якщо це так - типу підвищувального мікшування, яке застосовувалося. Метадані стану обробки підвищувального мікшування можуть бути корисними для реалізації знижувального мікшування (у постпроцесорі) у спадному напрямленні щодо декодера, наприклад, для зведення звукового вмісту програми таким чином, щоб воно було сумісним з одним з типів підвищувального мікшування (наприклад, Dolby Pro Logic або Dolby Pro Logic II Movie Mode, або Dolby Pro Logic II Music Mode, або Dolby Professional Upmixer), яке застосовувалося до програми. У варіантах здійснення, де кодований бітовий потік являє собою бітовий потік E-AC-3, метадані стану підвищувального мікшування можна використовувати в комбінації з іншими метаданими (наприклад, зі значенням поля «strmtyp» кадру) для визначення типу підвищувального мікшування (якщо воно мало місце), застосованого до каналу (каналів) програми. Значення поля «strmtyp» (у сегменті BSI кадру бітового потоку E-AC-3) вказує, належить звуковий вміст цього кадру незалежному потоку (що визначає програму) або незалежному вкладеному потоку (програми, яка містить кілька вкладених потоків або асоційована з ними), і тому може бути декодоване незалежно від будь-якого іншого вкладеного потоку, що вказується бітовим потоком E-AC-3, або того, чи належить звуковий вміст кадру залежному вкладеному потоку (програми, що містить кілька вкладених потоків або асоційованої

з ними), і тому воно повинне бути декодоване в комбінації з незалежним вкладеним потоком, з яким воно асоційоване; і

- метадані стану попередньої обробки даних, що служать ознакою того, чи виконувалася попередня обробка даних на звуковому вмісті кадру (перед кодуванням звукового вмісту в кодований бітовий потік, що генерується), і, якщо це так - типу виконаної попередньої обробки даних.

У деяких реалізаціях метадані стану обробки даних служать ознакою того:

- чи застосовувалося ослаблення навколишнього звуку (наприклад, чи послаблялися навколишні канали звукової програми на 3 дБ перед кодуванням),

- чи застосовувалося зрушення по фазі на 90 градусів (наприклад, до навколишніх каналів Ls і Rs звукової програми перед кодуванням),

- чи застосовувався фільтр пропущення нижніх частот до каналу LFE звукової програми перед кодуванням,

- чи відстежувався рівень каналу LFE програми в ході виробництва й, якщо це так, то який відстежений рівень каналу LFE щодо рівня широкосмугових звукових каналів програми,

- чи необхідно виконувати стиснення динамічного діапазону (наприклад, у декодері) на кожному блоці декодованого звукового вмісту програми, і, якщо це так, то який тип (і/або параметри) стиснення динамічного діапазону, що підлягає виконанню (наприклад, метадані стану попередньої обробки даних цього типу можуть служити ознакою того, який з типів профілю стиснення передбачався кодером для генерування контрольних значень стиснення динамічного діапазону, включених у кодований бітовий потік: Film Standard, Film Light, Music Standard, Music Light або Speech. У якості альтернативи, метадані стану попередньої обробки даних цього типу можуть вказувати, що на кожному кадрі декодованого звукового вмісту програми слід застосовувати інтенсивний стиснення динамічного діапазону (стиснення «comp») способом, обумовленим контрольними значеннями стиснення динамічного діапазону, включеними в кодований бітовий потік),

- чи була застосована обробка розтягування спектра й/або кодування зі зв'язуванням каналів для кодування конкретних діапазонів частот вмісту програми, і, якщо це так - які мінімальна й максимальна частоти частотних складових вмісту, на яких виконувалося розтягування спектра, і які мінімальна й максимальна частоти частотних складових вмісту, на яких виконувалося кодування зі зв'язуванням каналів. Відомості метаданих стану попередньої обробки даних цього типу можуть бути корисними при виконанні вирівнювання (у постпроцесорі) у спадному напрямку щодо декодера. Відомості як про зв'язування каналів, так і про розтягування спектра також корисні для оптимізації якості в ході операцій і застосувань перекодування. Наприклад, кодер може оптимізувати свою поведінку (у ході пристосовування етапів попередньої обробки даних, таких, як віртуалізація навушників, що підвищує мікшування і т.д.) на основі стану таких параметрів, як відомості про розтягування спектра й зв'язування каналів. Більше того, кодер міг би динамічно пристосовувати свої параметри зв'язування й розтягування спектра для відповідності й/або для оптимізації значень на основі стану вхідних (і аутентифікованих метаданих, і

- чи включені дані діапазону регулювання посилення діалогу в кодований бітовий потік, і, якщо це так - який доступний діапазон регулювання в ході виконання обробки посилення діалогу (наприклад, у заключній процесорі в спадному напрямку щодо декодера) для коректування рівня діалогового вмісту щодо рівня недіалогового вмісту звукової програми.

У деяких реалізаціях у корисне завантаження PIM кодованого бітового потоку, що підлягає виводу із кодера 100, (щаблем 107) включені метадані стану додаткової попередньої обробки даних (наприклад, метадані, що служать ознакою параметрів, що відносяться до навушників).

У деяких варіантах здійснення корисне завантаження LPSM, що включається (щаблем 107) у кадр кодованого бітового потоку (наприклад, бітового потоку E-AC-3, що служить ознакою щонайменше однієї звукової програми), містить LPSM у наступному форматі:

- заголовок (що містить, як правило, синхрослово, що ідентифікує початок корисного завантаження LPSM, після якого іде щонайменше одна величина-ідентифікатор, наприклад, значення версії формату LPSM, довжини, періоду, лічильника й асоціації вкладених потоків, зазначені нижче в Таблиці 2); і,

- після заголовка - щонайменше, одне значення покажчика діалогу (наприклад, параметр «Канал (канали) діалогу» за Таблицею 2), що вказує, вказують або не вказують відповідні аудіодані діалог (наприклад, які канали відповідних аудіоданих вказують діалог);

- щонайменше, одну величину відповідності гучності нормам (наприклад, параметр «Тип норм гучності» за Таблицею 2), що вказує чи відповідають відповідні аудіодані зазначеному набору норм гучності;

- щонайменше, одну величину обробки гучності (наприклад, один або кілька параметрів «Прапор виправлення стробованої гучності», «Тип виправлення гучності» за Таблицею 2), що вказує обробку гучності щонайменше одного типу, виконану на відповідних аудіоданих; і

5 - щонайменше, одну величину гучності (наприклад, один або кілька параметрів «Відносна стробована гучність ITU», «Стробована гучність мови ITU», «Короткострокова 3-секундна гучність ITU (EBU 3341)» і «Дійсне пікове значення» за Таблицею 2), що вказує щонайменше одну характеристику гучності (наприклад, пікову або середню гучність) відповідних аудіоданих.

У деяких варіантах здійснення кожний сегмент метаданих, що містить PIM і/або SSM (а також, необов'язково, інші метадані) містить: заголовок сегмента метаданих (а також, 10 необов'язково, додаткові базові елементи) і, після заголовка сегмента метаданих (або заголовка сегмента метаданих і інших базових елементів) щонайменше один сегмент корисного завантаження метаданих, що має наступний формат:

- заголовок корисного завантаження, як правило, що містить щонайменше одну величину-ідентифікатор (наприклад, значення версії формату SSM або PIM, довжини, періоду, лічильника 15 й асоціації вкладених потоків), і

- після заголовка корисного завантаження - SSM або PIM (або метадані іншого типу).

У деяких реалізаціях кожний із сегментів метаданих (іноді іменованих у цьому документі «контейнерами метаданих» або «контейнерами»), що вставляються шаблоном 107 у сегмент 20 зайвих бітів/поля ігнорованих даних (або в поле «addbsi», або в поле auxdata) кадру, має наступний формат:

- заголовок сегмента метаданих (що містить, як правило, синхрослово, що ідентифікує початок цього сегмента метаданих, за яким ідуть величини-ідентифікатори, наприклад, значення версії, довжини, періоду, лічильника елементів розширення й асоціації вкладених потоків, як зазначено нижче в Таблиці 1); і

25 - після заголовка сегмента метаданих - щонайменше, одна захисна величина (наприклад, значення згортки HMAC і контрольної суми аудіоданих), придатна для щонайменше однієї з дій: розшифрування, аутентифікації або перевірки вірогідності, щонайменше даних, що впливають: метаданих із цього сегмента метаданих або відповідних аудіоданих); і,

- також після заголовка сегмента метаданих - значення ідентифікатора («ID») і 30 конфігураційні значення корисного завантаження, що ідентифікують тип метаданих у кожному наступному корисному завантаженні метаданих, та що вказують щонайменше одну особливість конфігурації (наприклад, розмір) кожного такого корисного завантаження.

Кожне корисне завантаження метаданих іде за відповідним значенням ID корисного завантаження й конфігураційними значеннями корисного завантаження.

35 У деяких варіантах здійснення кожний із сегментів метаданих у сегменті зайвих бітів (або в полі auxdata, або в полі «addbsi») кадру має три рівні структури:

- структуру вищого рівня (наприклад, заголовок сегмента метаданих), що містить прапор, що вказує, чи містить метадані це поле зайвих бітів (або auxdata, або addbsi) щонайменше одне 40 значення ID, що вказує, метадані якого типу (типів) присутні, а також, як правило, величину, що вказує, скільки є присутніх бітів метаданих (наприклад, кожного типу) (якщо метадані присутні). Одним з типів метаданих, які можуть бути присутніми, є метадані PIM, іншим типом метаданих, які можуть бути присутніми, є метадані SSM, іншими типами метаданих, які можуть бути присутніми є метадані LPSM і/або метадані меж програми, і/або метадані досліджень в галузі засобів масової інформації;

45 - структуру проміжного рівня, що містить дані, асоційовані з кожним ідентифікованим типом метаданих (наприклад, значення заголовка корисного завантаження метаданих, захисних величин, ID корисного завантаження й конфігураційних значень корисного завантаження для кожного ідентифікованого типу метаданих); і

50 - структуру низового рівня, що містить корисне завантаження метаданих для кожного ідентифікованого типу метаданих (наприклад, послідовність значень PIM, якщо метадані PIM ідентифіковані як присутні, і/або значення метаданих іншого типу (наприклад, SSM або LPSM), якщо ці метадані іншого типу ідентифіковані як присутні).

Значення даних у такій трьохрівневій структурі можуть бути вкладеними. Наприклад, захисна величина (величини) для кожного корисного завантаження (наприклад, для кожного 55 корисного завантаження PIM або SSM, або інших метаданих), що ідентифікується структурами вищого й проміжного рівнів, може бути включена після корисного завантаження (і, таким чином, після заголовка корисного завантаження цього корисного завантаження метаданих), або захисна величина (величини) для всіх корисних завантажень метаданих, що ідентифікуються структурами вищого й проміжного рівнів, може бути включена після кінцевого корисного 60 завантаження в сегменті метаданих (і, таким чином, після заголовків корисних завантажень

метаданих усіх корисних завантажень цього сегмента метаданих).

В одному із прикладів (який буде описаний з посиланням на сегмент метаданих, або «контейнер», за фіг. 8), заголовок сегмента метаданих ідентифікує чотири корисні завантаження метаданих. Як показано на фіг. 8, цей заголовок сегмента метаданих містить синхрослово контейнера (що ідентифікується як «container sync») і значення версії й ID ключа. За заголовком сегмента метаданих слідує чотири корисні завантаження метаданих і біти захисту. За заголовком сегмента метаданих слідує значення ID корисного завантаження й конфігураційні значення корисного завантаження (наприклад, розмір корисного завантаження) для першого корисного завантаження (наприклад, для корисного завантаження PIM), а саме перше корисне завантаження іде за цим значенням ID і конфігураційними значеннями, за першим корисним завантаженням впливає значення ID корисного завантаження й конфігураційні значення корисного завантаження (наприклад, розмір корисного завантаження) для другого корисного завантаження (наприклад, для корисного завантаження SSM), а саме друге корисне завантаження іде за цим значенням ID і конфігураційними значеннями, за другим корисним завантаженням впливає значення ID корисного завантаження й конфігураційні значення корисного завантаження (наприклад, розмір корисного завантаження) для третього корисного завантаження (наприклад, для корисного завантаження LPSM), а саме третє корисне завантаження іде за цим значенням ID і конфігураційними значеннями, за третім корисним завантаженням впливає значення ID корисного завантаження й конфігураційні значення корисного завантаження (наприклад, розмір корисного завантаження) для четвертого корисного завантаження, а саме четверте корисне завантаження іде за цим значенням ID і конфігураційними значеннями, і за останнім корисним завантаженням впливає захисна величина (величини) (ідентифікована на фіг. 8 як «Дані захисту») для всіх або деяких зазначених корисних завантажень (або для структури вищого й проміжного рівнів, і для всіх або деяких корисних завантажень).

У деяких варіантах здійснення, якщо декодер 101 приймає бітовий аудіопотік, згенерований згідно з одним із варіантів здійснення винаходу зі значенням криптографічної хеш-функції, то декодер виконаний з можливістю синтаксичного аналізу й отримання цього значення криптографічної хеш-функції із блоку даних, визначеного з бітового потоку, при цьому зазначений блок містить метадані. Засіб 102 перевірки вірогідності може використовувати це значення криптографічної хеш-функції для перевірки вірогідності прийнятого бітового потоку і/або асоційованих метаданих. Наприклад, якщо засіб 102 перевірки вірогідності вважає метадані достовірними на підставі збігу між контрольним значенням криптографічної хеш-функції й значенням криптографічної хеш-функції, отриманим із цього блоку даних, то воно може скасовувати дію процесора 103 на відповідні аудіодані й викликати пропуск (незмінених) аудіоданих шаблоном 104 вибору. На додаток, необов'язково або в якості альтернативи, замість способу на основі значення криптографічної хеш-функції можна використовувати й інші криптографічні методи.

Кодек 100 за фіг. 2 може визначати (у відповідь на LPSM, а також, необов'язково, на метадані меж програми, отримані декодером 101), що модуль попередньої обробки/заключної обробки даних виконав (в елементах 105, 106 і 107) на аудіоданих, що підлягають кодуванню, обробку гучності будь-якого типу, і тоді може створювати (у генераторі 106) метадані стану обробки гучності, що містять конкретні параметри, використані та/або отримані при раніше виконаній обробці гучності. У деяких реалізаціях кодек 100 може створювати (і включати у виведений з нього кодований бітовий потік) метадані, що служать ознакою історії обробки даних на звуковому вмісті, оскільки кодек знає про типи обробки даних, виконаної на цьому звуковому вмісті.

Фіг. 3 являє собою блок-схему декодера (200), що представляє собою один з варіантів здійснення винахідницького модуля обробки аудіоданих, і пов'язаного з ним постпроцесора (300). Постпроцесор (300) також являє собою один з варіантів здійснення модуля обробки аудіоданих. Кожний з компонентів або елементів кодера 200 і постпроцесора 300 може бути виконаний як один або кілька процесів і/або одна або кілька схем (наприклад, мікросхем ASIC, матриць FPGA або інших інтегральних мікросхем), в апаратному забезпеченні, програмному забезпеченні або в комбінації апаратного й програмного забезпечення. Декодер 200 містить буфер 201 кадрів, синтаксичний аналізатор 205, аудіодекодер 202, щабель 203 перевірки вірогідності стану аудіоданих (засіб перевірки вірогідності) і щабель 204 генерування керуючих бітів, з'єднані так, як це показано. Як правило, декодер 200 також містить й інші елементи обробки даних (не показані).

Буфер 201 кадрів (буферна пам'ять) зберігає (наприклад, безстроковим чином) щонайменше один кадр кодованого бітового аудіопотоку, прийнятого декодером 200.

Послідовність кадрів кодованого бітового аудіопотоку направляється з буфера 201 у синтаксичний аналізатор 205.

Синтаксичний аналізатор 205 зв'язаний і виконаний з можливістю здобуття PIM і/або SSM (а також, необов'язково, інших метаданих, наприклад, LPSM) з кожного кадру кодованих вхідних аудіоданих для направлення щонайменше деяких із цих метаданих (наприклад, LPSM і метаданих меж програми, якщо які-небудь із них отримані, і/або PIM, і/або SSM) у засіб 203 перевірки вірогідності стану аудіоданих і на щабель 204 для направлення цих отриманих метаданих як виводу (наприклад, у постпроцесор 300), для здобуття аудіоданих з кодованих вхідних аудіоданих і для направлення отриманих аудіоданих у декодер 202.

Уведення кодованого бітового аудіопотоку в декодер 200 може являти собою один з наступних бітових потоків: бітовий потік AC-3, бітовий потік E-AC-3 або бітовий потік Dolby E.

Система за фіг. 3 також містить постпроцесор 300. Постпроцесор 300 містить буфер 301 кадрів і інші елементи обробки даних (не показані), у тому числі щонайменше один елемент обробки даних, пов'язаний з буфером 301. Буфер 301 кадрів зберігає (наприклад, безстроковим чином) щонайменше один кадр із декодованого бітового аудіопотоку, прийнятого постпроцесором з декодера 200. Елементи обробки даних постпроцесора 300 зв'язані й виконані з можливістю приймання й адаптивної обробки виводу послідовності кадрів декодованого бітового аудіопотоку з буфера 301 з використанням виводу метаданих з декодера 200 і/або виводу керуючих бітів із щаблем 204 декодера 200. Як правило, постпроцесор 300 виконаний з можливістю виконання адаптивної обробки на декодованих аудіоданих з використанням метаданих з декодера 200 (наприклад, адаптивної обробки гучності на декодованих аудіоданих з використанням значень LPSM, а також, необов'язково, метаданих меж програми, причому ця адаптивна обробка даних може бути заснована на стані обробки метаданих і/або на одній або декількох характеристиках аудіоданих, що вказуються LPSM для аудіоданих, що служать ознакою єдиної звукової програми).

Різні реалізації декодера 200 і постпроцесора 300 виконані з можливістю виконання різних варіантів здійснення способу за винаходом.

Аудіодекодер 202 декодера 200 виконаний з можливістю декодування аудіоданих, отриманих синтаксичним аналізатором 205, з метою генерування декодованих аудіоданих і для направлення цих декодованих аудіоданих у якості виводу (наприклад, у постпроцесор 300).

Засіб 203 перевірки вірогідності стану виконаний з можливістю аутентифікації й перевірки вірогідності метаданих, що направляються в нього. У деяких варіантах здійснення ці метадані являють собою блок даних (або містяться в блоці даних), який був включений у вхідний бітовий потік (наприклад, згідно з одним із варіантів здійснення справжнього винаходу). Цей блок може містити значення криптографічної хеш-функції (хеш-коду аутентифікації повідомлень, або «НМАС») для обробки метаданих і/або основаних на них аудіоданих (що доставляються в засіб 203 перевірки вірогідності із синтаксичного аналізатора 205 і/або декодера 202). У цих варіантах здійснення блок даних може містити цифровий підпис, тому модуль обробки аудіоданих у спадному напрямку може відносно легко аутентифікувати і перевіряти вірогідність зазначених метаданих стану обробки даних.

Для перевірки вірогідності метаданих (наприклад, у засобі 203 перевірки вірогідності) з метою забезпечення захищеної передачі й приймання метаданих і/або основаних на них аудіоданих, можна використовувати й інші криптографічні методи, у тому числі, без обмеження, будь-які з одного або декількох криптографічних методів, не заснованих на НМАС. Наприклад, перевірку вірогідності (з використанням такого криптографічного методу) можна виконувати в кожному модулі обробки аудіоданих, що приймає один з варіантів здійснення винахідницького бітового аудіопотоку з метою визначення того, чи підвергались метадані стану обробки гучності й відповідні аудіодані в бітовому потоці спеціальній обробці гучності (і/або є її результатом) (що вказується цими метаданими), і що вони не були модифіковані після виконання цієї спеціальної обробки гучності.

Для зазначення результатів операції перевірки вірогідності засіб 203 перевірки вірогідності стану направляє керуючі дані в генератор 204 керуючих бітів і/або направляє керуючі дані як вивід (наприклад, у постпроцесор 300). У відповідь на ці керуючі дані (а також, необов'язково, інші метадані, отримані із вхідного бітового потоку), щабель 204 може генерувати (і направляти в постпроцесор 300) одне з наступного:

- керуючі біти, що вказують, що вивід декодованих аудіоданих з декодера 202 піддавався обробці гучності конкретного типу (коли метадані LPSM указують, що вивід аудіоданих з декодера 202 піддавався обробці гучності конкретного типу, а керуючі біти із засобу 203 перевірки вірогідності вказують, що метадані LPSM є достовірними); або

- керуючі біти, що вказують, що вивід декодованих аудіоданих з декодера 202 слід піддавати

обробці гучності конкретного типу (наприклад, коли метадані LPSM указують, що вивід аудіоданих з декодера 202 не піддавався обробці гучності конкретного типу, або коли метадані LPSM указують, що вивід аудіоданих з декодера 202 піддавався обробці гучності конкретного типу, але керуючі біти із засобу 203 перевірки вірогідності вказують, що метадані LPSM не є достовірними).

У якості альтернативи, декодер 200 направляє метадані, отримані декодером 202 із вхідного бітового потоку, і метадані, отримані із вхідного бітового потоку синтаксичним аналізатором 205, у постпроцесор 300, і постпроцесор 300 виконує адаптивну обробку на декодованих аудіоданих, використовуючи ці метадані, або виконує перевірку вірогідності метаданих, а потім виконує адаптивну обробку на декодованих аудіоданих, використовуючи ці метадані, якщо перевірка вірогідності вказує, що ці метадані є достовірними.

У деяких варіантах здійснення, якщо декодер 200 приймає бітовий аудіопотік, згенерований згідно з одним із варіантів здійснення винаходу зі значенням криптографічної хеш-функції, то декодер виконаний з можливістю синтаксичного аналізу й здобуття цього значення криптографічної хеш-функції із блоку даних, визначеного з бітового потоку, при цьому зазначений блок містить метадані стану обробки гучності (LPSM). Засіб 203 перевірки вірогідності може використовувати це значення криптографічної хеш-функції для перевірки вірогідності прийнятого бітового потоку й/або асоційованих метаданих. Наприклад, якщо засіб 203 перевірки вірогідності знаходить метадані LPSM достовірними на основі збігу між контрольним значенням криптографічної хеш-функції й значенням криптографічної хеш-функції, отриманим із блоку даних, то воно може надавати сигнал модулю обробки аудіоданих у спадному напрямку (наприклад, постпроцесору 300, який може являти собою або містити модуль регулювання рівня гучності) пропуск (незмінених) аудіоданих бітового потоку. На додаток, необов'язково або в якості альтернативи, замість способу на основі значення криптографічної хеш-функції можна використовувати й інші криптографічні методи.

У деяких реалізаціях декодера 200 кодований бітовий потік, що приймається (і буферується у пам'яті 201) являє собою бітовий потік AC-3 або бітовий потік E-AC-3 і містить сегменти аудіоданих (наприклад, сегменти AB0-AB5 кадру, показаного на фіг. 4) і сегменти метаданих, де зазначені сегменти аудіоданих служать ознакою аудіоданих, і кожний з щонайменше деяких сегментів метаданих містить PIM або SSM (або інші метадані). Щабель 202 декодера (і/або синтаксичний аналізатор 205) виконано з можливістю здобуття метаданих з бітового потоку. Кожний із сегментів метаданих, що містять PIM і/або SSM (а також, необов'язково, інші метадані) включається в сегмент зайвих бітів кадру бітового потоку або в поле «addbsi» сегмента відомостей про бітовий потік («BSI») кадру бітового потоку, або в поле auxdata (наприклад, у сегмент AUX, показаний на фіг. 4) наприкінці кадру бітового потоку. Кадр бітового потоку може містити один або два сегменти метаданих, кожний з яких містить метадані, і, якщо цей кадр містить два сегменти метаданих, один з них може бути у полі addbsi кадру, а інший - у полі AUX кадру.

В деяких варіантах здійснення кожний сегмент метаданих (іноді іменований у цьому документі «контейнером») бітового потоку, який буферований в буфері 201, має формат, що включає заголовок сегмента метаданих (а також, необов'язково, інші обов'язкові, або «базові», елементи) і одне або кілька корисних завантажень метаданих, що слідує за заголовком сегмента метаданих. Метадані SIM, якщо вони присутні, містяться в одному з корисних завантажень метаданих (що ідентифікується за допомогою заголовка корисного завантаження та, як правило, що має формат першого типу). Метадані PIM, якщо вони присутні, містяться ще в одному корисному завантаженні метаданих (що ідентифікується за допомогою заголовка корисного завантаження та, як правило, має формат другого типу). Аналогічно, метадані кожного іншого типу (якщо вони присутні) містяться ще в одному з корисних завантажень (що ідентифікується за допомогою заголовка корисного завантаження та, як правило, має формат, специфічний для цього типу метаданих). Цей ілюстративний формат уможливлює зручний доступ до SSM, PIM і іншим метаданим в інші моменти часу, аніж під час декодування (наприклад, за допомогою постпроцесора 300 слідом за декодуванням або за допомогою процесора, виконаного з можливістю розпізнавання метаданих без виконання повного декодування на кодованому бітовому потоці), і уможливлює зручне й ефективне виявлення й виправлення помилок (наприклад, ідентифікації вкладених потоків) у ході декодування бітового потоку. Наприклад, під час відсутності доступу до SSM в ілюстративному форматі декодер 200 може невірно ідентифікувати вірну кількість вкладених потоків, асоційованих із програмою. Одне корисне завантаження метаданих у сегменті метаданих може містити SSM, ще одне корисне завантаження метаданих у сегменті метаданих може містити PIM, а також, необов'язково щонайменше ще одне корисне завантаження метаданих у сегменті метаданих

може містити інші метадані (наприклад, метадані стану обробки гучності, або «LPSM»).

У деяких варіантах здійснення корисне завантаження метаданих структури вкладених потоків (SSM), міститься в кадрі кодованого бітового потоку (наприклад, бітового потоку E-AC-3, що служить ознакою щонайменше однієї звукової програми), який буферований у буфері 201, містить метадані SSM у наступному форматі:

- заголовок корисного завантаження, як правило, що містить щонайменше одну величину-ідентифікатор (наприклад, 2-бітну величину, що служить ознакою версії формату SSM, а також, необов'язково значень довжини, періоду, лічильника й асоціації вкладених потоків); і,

- після заголовка - метадані незалежних вкладених потоків, що служать ознакою кількості незалежних вкладених потоків програми, що вказується цим бітовим потоком; і

- метадані залежних вкладених потоків, що служать ознакою того, чи містить кожний незалежний вкладений потік програми щонайменше один асоційований залежний вкладений потік (тобто, того, чи асоційований щонайменше один залежний вкладений потік із зазначеним кожним незалежним вкладеним потоком), і, якщо це так - кількості залежних вкладених потоків, асоційованих з кожним незалежним вкладеним потоком програми.

У деяких варіантах здійснення корисне завантаження метаданих відомостей про програму (PIM) міститься в кадрі кодованого бітового потоку (наприклад, кодованого бітового потоку E-AC-3, що служить ознакою щонайменше однієї звукової програми), який буферований у буфері 201, має наступний формат:

- заголовок корисного завантаження, як правило, що містить щонайменше одну величину-ідентифікатор (наприклад, величину, що служить ознакою версії формату PIM, а також, необов'язково значень довжини, періоду, лічильника й асоціації вкладених потоків); і,

- після заголовка - PIM у наступному форматі:

- метадані активних каналів, що служать ознакою кожного заглушеного каналу й кожного незаглушеного каналу звукової програми (тобто того, який канал (канали) програми містить звукову інформацію, а який (якщо він присутній) містить тільки мовчання (як правило, протягом часу певної тривалості)). У варіантах здійснення, де кодований бітовий потік являє собою бітовий потік AC-3 або E-AC-3, метадані активних каналів у кадрі бітового потоку можна використовувати в комбінації з додатковими метаданими бітового потоку (наприклад, з полем режиму звукового кодування («astmod») кадру й, якщо воно присутнє, з полем chapmap у цьому кадрі або в кадрі (кадрах) асоційованих вкладених потоків)) для визначення того, який канал (канали) програми містить звукову інформацію, а який містить мовчання;

- метадані стану обробки знижувального мікшування, які служать ознакою того, чи зазнала ця програма знижувального мікшування (перед або під час кодування), і, якщо це так - того, якого типу знижувальне мікшування застосовувалося. Метадані стану обробки знижувального мікшування можуть бути корисними для реалізації підвищувального мікшування (у постпроцесорі) у спадному напрямку щодо декодера, наприклад, для підвищувального мікшування звукового вмісту програми з використанням параметрів, які найбільше відповідають типу застосованого знижувального мікшування. У тих варіантах здійснення, де кодований бітовий потік являє собою бітовий потік AC-3 або E-AC-3, метадані стану обробки знижувального мікшування можна використовувати в комбінації з полем режиму звукового кодування («astmod») кадру для визначення типу знижувального мікшування (якщо воно мало місце), застосованого до каналу (каналів) програми;

- метадані стану обробки підвищувального мікшування, які служать ознакою того, чи зазнала програма підвищувального мікшування (наприклад, від меншої кількості каналів) перед або під час кодування, і, якщо це так - типу застосованого підвищувального мікшування. Метадані стану обробки підвищувального мікшування можуть бути корисними для реалізації знижувального мікшування (у постпроцесорі) у спадному напрямку щодо декодера, наприклад, для зведення звукового вмісту програми таким чином, щоб воно було сумісним з одним із типів підвищувального мікшування (наприклад, Dolby Pro Logic або Dolby Pro Logic II Movie Mode, або Dolby Pro Logic II Music Mode, або Dolby Professional Upmixer), яке застосовувалося до програми. У варіантах здійснення, де кодований бітовий потік являє собою бітовий потік E-AC-3, метадані стану підвищувального мікшування можна використовувати в комбінації з іншими метаданими (наприклад, зі значенням поля «strmtyp» кадру) для визначення типу підвищувального мікшування (якщо воно мало місце), застосованого до каналу (каналів) програми. Значення поля «strmtyp» (у сегменті BSI кадру бітового потоку E-AC-3) указує, належить звуковий вміст цього кадру незалежному потоку (що визначає програму) або незалежному вкладеному потоку (програми, яка містить кілька вкладених потоків або асоційована з ними), і тому може бути декодоване незалежно від будь-якого іншого вкладеного потоку, що вказується бітовим потоком E-AC-3, або того, чи належить звуковий вміст кадру

залежному вкладеному потоку (програми, що містить кілька вкладених потоків або асоційованої з ними), і тому воно повинне бути декодоване в комбінації з незалежним вкладеним потоком, з яким воно асоційоване; і

- метадані стану попередньої обробки даних, що служать ознакою того, чи виконувалася попередня обробка даних на звуковому вмісті кадру (перед кодуванням звукового вмісту в кодований бітовий потік, що генерується), і, якщо це так - типу виконаної попередньої обробки даних.

У деяких реалізаціях метадані стану обробки даних служать ознакою того:

- чи застосовувалося ослаблення навколишнього звуку (наприклад, чи послаблялися навколишні канали звукової програми на 3 дБ перед кодуванням),
 - чи застосовувалося зрушення по фазі на 90 градусів (наприклад, до навколишніх каналів Ls і Rs звукової програми перед кодуванням),
 - чи застосовувався фільтр пропущення нижніх частот до каналу LFE звукової програми перед кодуванням,
 - чи відстежувався рівень каналу LFE програми в ході виробництва й, якщо це так, то який відстежений рівень каналу LFE щодо рівня широкосмугових звукових каналів програми,
 - чи необхідно виконувати стиснення динамічного діапазону (наприклад, у декодері) на кожному блоці декодованого звукового вмісту програми, і, якщо це так, то який тип (і/або параметри) стиснення динамічного діапазону, що підлягає виконанню (наприклад, метадані стану попередньої обробки даних цього типу можуть служити ознакою того, який з типів профілю стиснення передбачався кодером для генерування контрольних значень стиснення динамічного діапазону, що містяться у кодованому бітовому потоці: Film Standard, Film Light, Music Standard, Music Light або Speech. У якості альтернативи, метадані стану попередньої обробки даних цього типу можуть указувати, що на кожному кадрі декодованого звукового вмісту програми слід застосовувати інтенсивне стиснення динамічного діапазону (стиснення «comp») способом, обумовленим контрольними значеннями стиснення динамічного діапазону, включеними в кодований бітовий потік),
 - чи застосовувалася обробка розтягування спектра й/або кодування зі зв'язуванням каналів для кодування конкретних діапазонів частот вмісту програми, і, якщо це так - які мінімальна й максимальна частоти частотних складових вмісту, на яких виконувалося розтягування спектра, і які мінімальна й максимальна частоти частотних складових вмісту, на яких виконувалося кодування зі зв'язуванням каналів. Відомості метаданих стану попередньої обробки даних цього типу можуть бути корисними при виконанні вирівнювання (у постпроцесорі) у спадному напрямку щодо декодера. Відомості як про зв'язування каналів, так і про розтягування спектра також корисні для оптимізації якості в ході операцій і застосувань перекодування. Наприклад, кодер може оптимізувати свою поведінку (у ході пристосовування етапів попередньої обробки даних, таких, як віртуалізація навушників, що підвищує міксування і т.д.) на основі стану таких параметрів, як відомості про розтягування спектра й зв'язування каналів. Більше того, кодер міг би динамічно пристосовувати свої параметри зв'язування й розтягування спектра для відповідності й/або для оптимізації значень на основі стану вхідних (і аутентифікованих) метаданих, і
 - чи включені дані діапазону регулювання посилення діалогу в кодований бітовий потік, і, якщо це так - який доступний діапазон регулювання в ході виконання обробки посилення діалогу (наприклад, у постпроцесорі в спадному напрямку щодо декодера) для коректування рівня діалогового вмісту щодо рівня недіалогового вмісту звукової програми.
- У деяких варіантах здійснення корисне завантаження LPSM, що включається в кадр кодованого бітового потоку (наприклад, бітового потоку E-AC-3, що служить ознакою щонайменше однієї звукової програми), який буферований у буфері 201, містить LPSM у наступному форматі:
- заголовок (що містить, як правило, синхрослово, що ідентифікує початок корисного завантаження LPSM, за яким іде щонайменше одна величина-ідентифікатор, наприклад, значення версії формату LPSM, довжини, періоду, лічильника й асоціації вкладених потоків, зазначені нижче в Таблиці 2); і,
 - після заголовка - щонайменше, одне значення покажчика діалогу (наприклад, параметр «Канал (канали) діалогу» за Таблицею 2), що вказує, указують або не вказують діалог відповідні аудіодані (наприклад, які канали відповідних аудіоданих указують діалог);
 - щонайменше, одну величину відповідності гучності нормам (наприклад, параметр «Тип норм гучності» за Таблицею 2), що вказує чи відповідають відповідні аудіодані зазначеному набору норм гучності;
 - щонайменше одну величину обробки гучності (наприклад, один або кілька параметрів

«Прапор виправлення стробованої гучності», «Тип виправлення гучності» за Таблицею 2) що вказує обробку гучності щонайменше одного типу, виконану на відповідних аудіоданих; і

- щонайменше, одну величину гучності (наприклад, один або кілька параметрів «Відносна стробована гучність ITU», «Стробована гучність мови ITU», «Короткострокова 3-секундна гучність ITU (EBU 3341)» і «Дійсне пікове значення» за Таблицею 2), що вказує щонайменше одну характеристику гучності (наприклад, пікову або середню гучність) відповідних аудіоданих.

У деяких реалізаціях синтаксичний аналізатор 205 (і/або щабель 202 декодера) виконаний з можливістю здобуття із сегмента зайвих бітів або з поля «addbsi», або з поля auxdata кадру бітового потоку кожного сегмента метаданих, що має наступний формат:

- заголовок сегмента метаданих (що містить, як правило, синхрослово, що ідентифікує початок, сегмента метаданих, за яким іде щонайменше одна величина-ідентифікатор, наприклад, значення версії, довжини, періоду, лічильника елементів розширення й асоціації вкладених потоків); і

- після заголовка сегмента метаданих - щонайменше, одна захисна величина (наприклад, значення згортки HMAC і контрольної суми аудіоданих за Таблицею 1), придатна для щонайменше розшифрування, аутентифікації або перевірки вірогідності, щонайменше даних, що впливають: метаданих із сегмента метаданих або відповідних аудіоданих; і

- також після заголовка сегмента метаданих - значення ідентифікатора корисного завантаження метаданих («ID») і конфігураційні значення корисного завантаження, що ідентифікують тип і щонайменше одну особливість конфігурації (наприклад, розмір) кожного наступного корисного завантаження метаданих.

Кожний сегмент корисного завантаження метаданих (переважно вищевказаний формат, що має) іде за відповідним ідентифікатором ID корисного завантаження метаданих і конфігураційними значеннями корисного завантаження.

У більш загальному значенні, кодований бітовий аудіопотік, згенерований кращими варіантами здійснення винаходу, має структуру, що забезпечує механізм розмітки елементів метаданих і вкладених елементів як базових (обов'язкових) або (необов'язкових) елементів розширення, або додаткових елементів. Це дозволяє масштабувати швидкість передачі даних бітового потоку (у тому числі його метаданих) на безліч додатків. Базові (обов'язкові) елементи кращого синтаксису бітового потоку також повинні бути здатні сигналізувати про те, що (необов'язкові) елементи розширення присутні (усередині смуги) і/або перебувають у віддаленому місці розташування (поза смугою).

Присутність базового (базових) елемента (елементів) необхідна в кожному кадрі бітового потоку. Деякі додаткові елементи базових елементів є необов'язковими й можуть бути присутніми у будь-якій комбінації. Присутність елементів розширення в кожному кадрі не є необхідним (для обмеження невігідних витрат бітової швидкості передачі даних). Таким чином, елементи розширення в деяких кадрах можуть бути присутніми, а в інших - ні. Деякі додаткові елементи елемента розширення є необов'язковими й можуть бути присутніми у будь-якій комбінації, у той час як деякі додаткові елементи елемента розширення можуть бути обов'язковими (тобто якщо елемент розширення присутній в кадрі бітового потоку).

В одному із класів варіантів здійснення генерується кодований бітовий аудіопотік (наприклад, модулем обробки аудіоданих, що втілюють винахід), що містить послідовність сегментів аудіоданих і сегментів метаданих. Сегменти аудіоданих служать ознаками аудіоданих, кожний з щонайменше деяких із сегментів метаданих містить PIM і/або SSM (а також, необов'язково, метадані щонайменше ще одного типу), і зазначені сегменти аудіоданих ущільнені з тимчасовим поділом із сегментами метаданих. В переважних варіантах здійснення винаходу в цьому класі кожний із сегментів метаданих має кращий формат, що описується у цьому документі.

В одному із переважних форматів кодований бітовий потік являє собою бітовий потік AC-3 або бітовий потік E-AC-3, і кожний сегмент метаданих, що містить SSM і/або PIM, включений (наприклад, шаблоном 107 однієї із переважних реалізацій кодера 100) у якості додаткових відомостей про бітовий потік у поле «addbsi» (показане на фіг. 6) сегмента відомостей про бітовий потік («BSI») кадру бітового потоку або в поле auxdata кадру бітового потоку, або в сегмент зайвих бітів кадру бітового потоку.

У цьому переважному форматі кожний з кадрів містить сегмент метаданих (іноді іменований контейнером метаданих, або контейнером) у сегменті зайвих бітів (або в полі addbsi) кадру. Сегменти метаданих містять обов'язкові елементи (спільно іменовані «базовим елементом»), показані нижче в Таблиці 1 (і можуть містити необов'язкові елементи, показані в Таблиці 1). Щонайменше, деякі з необхідних елементів, показаних у Таблиці 1, містяться в заголовку сегмента метаданих, але деякі можуть бути включені й в іншому місці сегмента метаданих.

Таблиця 1

Параметр	Опис	Обов'язковий (М)/ необов'язковий (О)
		М
SYNC [ID]		М
Версія базового елемента		М
Довжина базового елемента		М
Період базового елемента (xxx)		М
Лічильник елементів розширення	Указує кількість елементів розширення метаданих, асоційованих з базовим елементом. Це значення може надавати збільшення/негативне збільшення в міру проходження бітового потоку від виробництва через поширення до остаточного випуску	М
Асоціація вкладених потоків	Описує вкладений потік (потоки), з якими асоційований базовий елемент	М
Сигнатура (згортка HMAC)	256-бітна згортка HMAC (з використанням алгоритму SHA-2), обчислена згідно з аудіоданими, базовим елементом й усім елементам розширення всього кадру	М
Зворотний відлік меж PGM	Поле з'являється тільки для деякої кількості кадрів у голові або у хвості файлу/потоків звукової програми. Тому для сигналізації включення цього параметра можна використовувати зміну версії базового елемента.	О
Контрольна сума аудіоданих	Контрольна сума аудіоданих, взята згідно з деякою кількістю дискретних значень PCM аудіоданих, що представляються полем періоду базового елемента.	О
Контрольна сума відеоданих	Контрольна сума відеоданих, по взята згідно з деякою кількістю стислих дискретних значень відеоданих (якщо вони присутні), що представляються полем періоду базового елемента.	О
URL/UUID	Це поле визначено для переносу ідентифікатора URL і/або UUID (воно може бути надлишковим стосовно контрольної суми), який посилається на зовнішнє місце розташування додаткового вмісту, програми (сутності) і/або на метадані, асоційовані з бітовим потоком.	О

У цьому переважному форматі кожний сегмент метаданих (у сегменті зайвих бітів або в полі addbsi або auxdata кадру кодованого бітового потоку), що містить SSM, PIM або LPSM, містить заголовок сегмента метаданих (а також, необов'язково, додаткові базові елементи), і, після цього заголовка сегмента метаданих (або цього заголовка сегмента метаданих і додаткових базових елементів), одну або кілька корисних завантажень метаданих. Кожне корисне завантаження метаданих містить заголовок корисного завантаження метаданих (що вказує конкретний тип метаданих (наприклад, SSM, PIM або LPSM), включених у це корисне завантаження), за яким ідуть метадані даного конкретного типу. Як правило, заголовок корисного завантаження метаданих містить наступні величини (параметри):

- ідентифікатор ID корисного завантаження (ідентифікуючий тип метаданих, наприклад, SSM, PIM або LPSM), що слідує за заголовком сегмента метаданих (який може містити величини, зазначені в Таблиці 1);

- конфігураційне значення корисного завантаження (як правило, що вказує розмір цього корисного завантаження), що слідує за ID корисного завантаження; а також, необов'язково,
 - додаткові конфігураційні значення корисного завантаження (наприклад, значення зміщення, що вказує кількість кадрів дискретних значень аудіоданих від початку кадру до першого дискретного значення, якому належить це корисне завантаження, і значення пріоритету корисного завантаження, наприклад, що вказує умову, за якої це корисне завантаження може бути скасовано).

Як правило, метадані із зазначеного корисного завантаження мають один з наступних форматів:

- метадані корисного завантаження, що представляють собою метадані SSM, містять метадані незалежних вкладених потоків, що служать ознакою кількості незалежних вкладених потоків у програмі, що вказується цим бітовим потоком; і метадані залежних вкладених потоків, що служать ознакою того, чи містить кожний незалежний вкладений потік програми щонайменше один асоційований з ним залежний вкладений потік, і, якщо це так - кількості залежних вкладених потоків, асоційованих з кожним незалежним вкладеним потоком цієї програми;

- метадані корисного завантаження, що представляють собою метадані PIM, містять метадані активних каналів, що служать ознакою того, який канал (канали) звукової програми містить звукову інформацію, а який (якщо він присутній) - містить тільки мовчання (як правило, протягом тривалості кадру); метадані стану обробки знижувального міксування, що служать ознакою того, чи зазнала ця програма знижувального міксування (перед кодуванням або в ході його), і, якщо це так - типу застосованого знижувального міксування; метадані стану підвищувального міксування, що служать ознакою того, чи зазнала ця програма підвищувального міксування (наприклад, від меншої кількості каналів) перед або під час кодування, і, якщо це так - типу застосованого підвищувального міксування; і метадані стану попередньої обробки даних, що служать ознакою того, чи виконувалася на звуковому вмісті кадру попередня обробка даних (перед кодуванням звукового вмісту в кодований бітовий потік, що генерується), і, якщо це так - типу виконаної попередньої обробки даних; або

- метадані корисного завантаження являють собою метадані LPSM, що мають формат, зазначений в наступній таблиці (Таблиця 2).

Таблиця 2

Параметр LPSM [інтелектуальна гучність]	Опис	Кількість унікальних станів	Обов'язковий (М)/необов'язковий (О)	Частота вставки (період відновлення параметра)
Версія LPSM			М	
Період LPSM (xxx)	Застосуємо тільки до полів xxx		М	
Лічильник LPSM			М	
Асоціація вкладених потоків LPSM			М	
Канал (канали) діалогу	Указує, яка комбінація звукових каналів L, C і R містить мовлення протягом попередніх 0,5 секунд. Якщо мовлення відсутнє в жодній з комбінацій L, C або R, то цей параметр буде вказувати «No dialog» (відсутність діалогу)	8	М	~0,5 секунд (як правило)
Тип норм гучності	Указує, що асоційований потік аудіоданих відповідає конкретному набору норм (наприклад, ATSC A/85 або EBU R128)	8	М	Кадр

Таблиця 2

Параметр LPSM [інтелектуальна гучність]	Опис	Кількість унікальних станів	Обов'язковий (М)/необов'язковий (О)	Частота вставки (період відновлення параметра)
Прапор виправлення стробованої гучності діалогу	Указує, чи був асоційований аудіопотік виправлений на основі стробованого діалогу	2	О (є присутнім тільки у випадку, якщо Loudness_Regulation_ Type указує, що відповідні аудіодані є НЕВИПРАВЛЕНИМИ)	Кадр
Тип виправлення метаданих	Указує, чи був асоційований аудіопотік виправлений контролером гучності й динамічного діапазону з нескінченим випереджувальним переглядом (на основі файлу) або в реальному часі (RT).	2	О (є присутнім тільки у випадку, якщо Loudness_Regulation_ Type указує, що відповідні аудіодані є НЕВИПРАВЛЕНИМИ)	Кадр
Відносна стробована гучність ITU (INF)	Указує сумарну гучність згідно з ITU-R BS.1770-3 для асоційованого аудіопотоку без застосування метаданих (наприклад, 7 бітів: - 58→+5,5 одиниць гучності, зваженої за кривою К, щодо повної шкали, із кроком 0,5 одиниць гучності, зваженої за кривою К, щодо повної шкали).	128	О	1 с
Стробована гучність мови ITU (INF)	Указує сумарну гучність згідно з ITU-R BS.1770-3 для асоційованого аудіопотоку без застосування метаданих (наприклад, 7 бітів: - 58→+5,5 одиниць гучності, зваженої за кривою К, щодо повної шкали, із кроком 0,5 одиниць гучності, зваженої за кривою К, щодо повної шкали).	128	О	1 с
Короткострокова 3-секундна гучність ITU (EBU 3341)	Указує 3-секундну нестробовану гучність згідно з ITU (ITU-R BS.1771-1) для асоційованого аудіопотоку без застосування метаданих (вікно пакетної передачі змінної тривалості) при частоті вставки ~10 Гц (наприклад, 8 бітів: 116→+11,5 одиниць гучності, зваженої за кривою К, щодо повної шкали, із кроком 0,5 одиниць гучності, зваженої за кривою К, щодо повної шкали).	256	О	0,1 с

Таблиця 2

Параметр LPSM [інтелектуальна гучність]	Опис	Кількість унікальних станів	Обов'язковий (М)/необов'язковий (О)	Частота вставки (період відновлення параметра)
Дійсне пікове значення	Указує дійсне пікове значення (дБ TP) згідно з Додатком 2 ITU-R BS.1770-3 для асоційованого аудіопотоку без застосування метаданих (тобто найбільше значення за весь період кадру, що сигналізується у поле періоду елемента) 116→+11,5 одиниць гучності, зваженої за кривою K, щодо повної шкали, із кроком 0,5 одиниць гучності, зваженої за кривою K, щодо повної шкали.	256	О	0,5 с
Зсув знижувального мікшування	Указує зсув знижувального мікшування гучності			
Межа програми	Указує у кадрах, коли зустрінеться або зустрілася межа програми. Коли межа програми не є межею кадру, необов'язкове зміщення дискретних значень буде вказувати, як далеко в кадрі зустрічається фактична межа програми			

В іншому переважному форматі кодованого бітового потоку, згенерованого відповідно до винаходу, бітовий потік являє собою бітовий потік AC-3 або E-AC-3, і кожний із сегментів метаданих, що включає PIM і/або SSM (а також, необов'язково, метадані щонайменше ще одного типу) включений (наприклад, шаблоном 107 кращої реалізації кодера 100) у будь-який із наступних сегментів: сегмент зайвих бітів кадру бітового потоку; або в поле «addbsi» (показане на фіг. 6) сегмента відомостей про бітовий потік («BSI») кадру бітового потоку; або в поле auxdata (наприклад, у сегмент AUX, показаний на фіг. 4) наприкінці кадру бітового потоку. Кадр може містити один або два сегменти метаданих, кожний з яких містить PIM і/або SSM, і, у деяких варіантах здійснення, якщо кадр містить два сегменти метаданих, один може бути присутнім у полі addbsi кадру, а інший - у полі AUX кадру. Кожний сегмент метаданих переважно має формат, зазначений вище з посиланням на наведену вище Таблицю 1 (тобто, він містить базові елементи, зазначені в Таблиці 1, за якими ідуть ID корисних завантажень (ідентифікуючі тип метаданих у кожному корисному завантаженні цього сегмента метаданих), конфігураційні значення корисних завантажень і кожна з корисних завантажень). Кожний сегмент метаданих, що містить LPSM, переважно має формат, зазначений вище з посиланнями на наведені вище Таблиці 1 і 2 (тобто, він містить базові елементи, зазначені в Таблиці 1, за якими іде ID корисного завантаження (ідентифікуючий метадані як LPSM) і конфігураційні значення корисного завантаження, за якими іде корисне завантаження (дані LPSM, що мають формат, зазначений у Таблиці 2)).

В іншому переважному форматі кодований бітовий потік являє собою бітовий потік Dolby E, і кожний із сегментів метаданих, що містить PIM і/або SSM (а також, необов'язково, інші метадані) являє собою перші N місць розташування дискретних значень інтервалу захисної смуги Dolby E. Бітовий потік Dolby E, що містить такий сегмент метаданих, який включає LPSM, як правило, містить величину, що служить ознакою довжини корисного завантаження LPSM, яка

сигналізується у слові Pd преамбули SMPTE 337M (частота повторення слова Pa згідно з SMPTE 337M переважно залишається ідентичною частоті кадрів асоційованого відеозображення).

У переважному форматі, де кодований бітовий потік являє собою бітовий потік E-AC-3, кожний із сегментів метаданих, що містить PIM і/або SSM (а також, необов'язково, LPSM і/або інші метадані) включений (щаблем 107 кращої реалізації кодера 100) у якості додаткових відомостей про бітовий потік у сегмент зайвих бітів або в поле «addbsi» сегмента відомостей про бітовий потік («BSI») кадру бітового потоку. Нижче ми опишемо додаткові особливості кодування бітового потоку E-AC-3 з LPSM у цьому переважному форматі:

1. у ході генерування бітового потоку E-AC-3, тоді як кодер E-AC-3 (що вставляє значення LPSM у бітовий потік) є «активним», бітовий потік для кожного кадру (синхрокадру), що генерується, повинен містити блок метаданих (що містить LPSM), який переноситься у полі addbsi (або в сегменті зайвих бітів) цього кадру. Біти, необхідні для переносу цього блоку метаданих не повинні підвищувати бітову швидкість передачі даних кодера (довжину кадру);

2. кожний блок метаданих (що містить LPSM) повинен містити наступну інформацію:

- прапор `loudness_correction_type_flag`: де «1» указує, що гучність відповідних аудіоданих була виправлена у висхідному напрямку відносно кодера, і «0» указує, що гучність була виправлена засобом виправлення гучності, впровадженим у кодер (наприклад, у процесор 103 гучності кодера 100 за фіг. 2);

- `speech_channel`: указує, який канал (канали) джерела містять мовлення (протягом попередніх 0,5 с). Якщо мовлення не виявлено, то це може бути так і зазначено;

- `speech_loudness`: указує інтегральну гучність мовлення кожного відповідного звукового каналу, що містить мовлення (протягом попередніх 0,5 с);

- `ITU_loudness`: указує інтегральну гучність згідно з ITU BS.1770-3 кожного відповідного звукового каналу; і

- `gain`: складений коефіцієнт (коефіцієнти) посилення гучності для обігу в декодері (з метою демонстрації оборотності);

3. Тоді як кодер E-AC-3 (що вставляє значення LPSM у бітовий потік) є «активним» і приймає кадр AC-3 із прапором «trust», контролер гучності кодера (наприклад, процесор 103 гучності кодера 100 за фіг. 2), слід обійти. Значення `dialnorm` і `DRC` джерела, «що заслуговує довіри», повинні бути пропущені (наприклад, генератором 106 кодера 100) у компонент кодера E-AC-3 (наприклад, на щабель 107 кодера 100). Генерування блоків LPSM триває, і прапор `loudness_correction_type_flag` устанавлюється на «1». Послідовність обходу контролера гучності повинна бути синхронізована з початком того декодованого кадру AC-3, у якому з'являється прапор «trust». Послідовність обходу контролера гучності повинна бути виконана в такий спосіб: керуючий елемент `leveler_amount` зазнає негативного збільшення від значення 9 до значення 0 протягом 10 періодів аудіоблока (тобто 53,3 мс), і управляючий елемент `leveler_back_end_meter` міститься в режимі обходу (ця операція повинна в результаті призводити до безрозривного переходу). Термін обхід регулятора рівня, «що заслуговує довіри», передбачає те, що значення `dialnorm` джерела бітового потоку повторно використовується також і як вивід кодера (наприклад, якщо бітовий потік з джерела, «що заслуговує довіри», має значення `dialnorm` -30, то вивід повинен використовувати -30 для вихідного значення `dialnorm`);

4. Тоді як кодер E-AC-3 (що вставляє значення LPSM у бітовий потік) є «активним» і приймає кадр AC-3 без прапора «trust», впроваджений у кодер контролер гучності (наприклад, процесор 103 кодера 100 за фіг. 2) повинен бути активним. Генерування блоків LPSM триває, і прапор `loudness_correction_type_flag` устанавлюється на «0». Послідовність активації контролера гучності повинна бути синхронізована з початком того кодованого кадру AC-3, у якому зникає прапор «trust». Послідовність активації контролера гучності повинна бути виконана в такий спосіб: керуючий елемент `leveler_amount` зазнає збільшення від значення 0 до значення 9 протягом 1 періоду аудіоблока (тобто 5,3 мс), і управляючий елемент `leveler_back_end_meter` міститься в режимі «активний» (ця операція повинна в результаті призводити до безрозривного переходу й включати відміну інтеграції `back_end_meter`); і

5. під час кодування графічний користувацький інтерфейс (GUI) повинен указувати користувачеві наступні параметри: «InputAudioProgram:[Trusted/Untrusted]» (Вхідна звукова програма [заслуговує/не заслуговує довіри]) - стан цього параметра базується на присутності прапора «trust» у вхідному сигналі; і «Real-time Loudness Correction: [Enabled/Disabled]» (Виправлення гучності в реальному часі: [Включене/Виключене]) - стан цього параметра базується на тому, чи є активним цей контролер гучності.

У процесі декодування бітового потоку AC-3 або E-AC-3, що містить метадані LPSM (у переважному форматі), включені в сегмент зайвих бітів, або в поле ігнорованих даних, або в

поле «addbsi» сегмента відомостей про бітовий потік («BSI») кожного кадру бітового потоку, декодер повинен виконувати синтаксичний аналіз блоку даних LPSM (у сегменті зайвих бітів або полі field) і передавати всі отримані значення LPSM у графічний користувацький інтерфейс (GUI). Цей набір отриманих значень LPSM обновляється для кожного кадру.

5 В іншому переважному форматі кодованого бітового потоку, згенерованого відповідно до винаходу, цей кодований бітовий потік являє собою бітовий потік AC-3 або бітовий потік E-AC-3, і кожний із сегментів метаданих, що містить PIM і/або SSM (а також, необов'язково, LPSM і/або інші метадані), включений (наприклад, шаблем 107 однієї із переважних реалізацій кодера 100) у сегмент зайвих бітів або в сегмент Aux, або, як додаткові відомості про бітовий потік, у поле «addbsi» (показане на фіг. 6) сегмента відомостей про бітовий потік («BSI») кадру бітового потоку. У цьому форматі (який являє собою зміну вищепри описаного формату з посиланнями на Таблиці 1 і 2) кожне з полів addbsi (або Aux, або зайвих бітів), що містить LPSM, містить наступні значення LPSM:

15 - базові елементи, зазначені в Таблиці 1, за якими іде ідентифікатор ID корисного завантаження (що ідентифікує метадані як LPSM) і конфігураційні значення корисного завантаження, за якими іде корисне завантаження (дані LPSM), що має наступний формат (подібний з обов'язковими елементами, зазначеними вище в Таблиці 2):

-- версія корисного завантаження LPSM: 2-бітне поле, що вказує версію корисного завантаження LPSM;

20 -- поле dialchan: 3-бітне поле, що вказує, чи містять лівий, правий і/або центральний канали відповідних аудіоданих діалог, що проговорюється. Розподіл бітів поля dialchan може бути наступним: біт 0, що вказує присутність діалогу в лівому каналі, зберігається в самому старшому біті поля dialchan; і біт 2, що вказує присутність діалогу в центральному каналі, зберігається в самому молодшому біті поля dialchan. Кожний біт поля dialchan встановлений на «1», якщо відповідний канал містить виражений словами діалог протягом попередніх 0,5 секунд програми;

-- поле loudregtyp: 4-бітне поле, що вказує, якому стандарту норм гучності відповідає гучність програми. Установка loudregtyp на «000» вказує, що метадані LPSM не вказують відповідність нормам гучності. Наприклад, одне значення цього поля (наприклад, 0000) може вказувати, що відповідність стандарту норм гучності не зазначене, інше значення цього поля (наприклад, 0001) може вказувати, що аудіодані програми відповідають стандарту ATSC A/85, а ще одне значення цього поля (наприклад, 0010) може вказувати, що аудіодані цієї програми відповідають стандарту EBU R128. У цьому прикладі, якщо полю привласнене будь-яке значення, крім «0000», то в корисному завантаженні за ним повинні йти поля loudcorrldialgat і loudcorrtyp;

35 -- поле loudcorrldialgat: однобітне поле, що вказує, чи було застосоване виправлення стробованої гучності діалогу. Якщо гучність програми була виправлена з використанням стробування діалогу, значення loudcorrldialgat встановлюється на «1». Інакше воно встановлюється на «0»;

40 -- поле loudcorrtyp: однобітне поле, що вказує тип виправлення гучності, застосованого до програми. Якщо гучність програми була виправлена в процесі виправлення гучності з нескінченним випереджувальним переглядом, то значення поля loudcorrtyp встановлюється на «0». Якщо гучність програми була виправлена з використанням комбінації виміру гучності в реальному часі й керування динамічним діапазоном, то значення цього поля встановлюється на «1»;

45 -- поле loudrelgate: однобітне поле, що вказує, чи існують дані відносної стробованої гучності (ITU). Якщо поле loudrelgate встановлене на «1», то в корисному завантаженні за ним повинне йти 7-бітне поле ituloudrelgat;

50 -- поле loudrelgat: 7-бітне поле, що вказує відносну стробовану гучність програми (ITU). Це поле вказує інтегральну гучність звукової програми, виміряну у відповідності зі стандартом ITU-R BS.1770-3 без яких-небудь регулювань коефіцієнтів посилення через застосування dialnorm і стиснення динамічного діапазону (DRC). Значення 0-127 інтерпретують як інтервал від -58 одиниць гучності, зваженої за кривою K, щодо повної шкали, до +5,5 одиниць гучності, зваженої за кривою K, щодо повної шкали, із кроком 0,5 одиниць гучності, зваженої за кривою K, щодо повної шкали;

55 -- поле loudspchgate: однобітне поле, що вказує, чи існують дані стробованої гучності мови (ITU). Якщо поле loudspchgate встановлене на «1», то в корисному завантаженні за ним повинне йти 7-бітне поле loudspchgat;

60 -- поле loudspchgat: 7-бітне поле, що вказує гучність програми зі стробованим мовленням. Це поле вказує інтегральну гучність усієї відповідної звукової програми, виміряну відповідно до

формули (2) стандарту ITU-R BS.1770-3 і без будь-яких регулювань посилення через застосування `diapnorm` і стиснення динамічного діапазону. Значення 0-127 інтерпретують як інтервал -58-+5,5 одиниць гучності, зваженої за кривою K, щодо повної шкали, із кроком 0,5 одиниць гучності, зваженої за кривою K, щодо повної шкали;

5 -- поле `loudstrm3se`: однобітне поле, що вказує, чи існують короткострокові (3-секундні) дані гучності. Якщо це поле встановлене на «1», то в корисному завантаженні за ним повинне впливати поле `loudstrm3s`;

10 -- поле `loudstrm3s`: 7-бітне поле, що вказує нестробовану гучність попередніх 3 секунд відповідної звукової програми, виміряну у відповідності зі стандартом ITU-R BS.1771-1 і без будь-яких регулювань посилення через застосування `diapnorm` і стиснення динамічного діапазону. Значення 0-256 інтерпретують як інтервал від -116 одиниць гучності, зваженої за кривою K, щодо повної шкали, до +11,5 одиниць гучності, зваженої за кривою K, щодо повної шкали, із кроком 0,5 одиниць гучності, зваженої за кривою K, щодо повної шкали;

15 -- поле `truepk`: однобітне поле, що вказує, чи існують дані дійсної пікової гучності. Якщо поле `truepk` встановлене на «1», то в корисному завантаженні за ним повинне йти 8-бітне поле `truepk`; і

20 -- поле `truepk`: 8-бітне поле, що вказує опорне значення дійсного піка програми, виміряне відповідно до Додатка 2 стандарту ITU-R BS.1770-3 і без будь-яких регулювань посилення через застосування `diapnorm` і стиснення динамічного діапазону. Значення 0-256 інтерпретують як інтервал від -116 одиниць гучності, зваженої за кривою K, щодо повної шкали, до +11,5 одиниць гучності, зваженої за кривою K, щодо повної шкали, із кроком 0,5 одиниць гучності, зваженої за кривою K, щодо повної шкали.

25 У деяких варіантах здійснення базовий елемент сегмента метаданих у сегменті зайвих бітів або в полі `auxdata` (або `"addbsi"`) кадру бітового потоку AC-3 або бітового потоку E-AC-3 містить заголовок сегмента метаданих (як правило, що містить величини-ідентифікатори, наприклад, версії) і, після цього заголовка сегмента метаданих: величини, що служать ознаками того, чи включені для цього сегмента метаданих дані контрольної суми (або інші захисні величини), величини, що служать ознаками того, чи існують зовнішні дані (що відносяться до аудіоданих відповідно до метаданих цього сегмента метаданих), значення ID корисного завантаження й конфігураційні значення для метаданих кожного типу (наприклад, PIM і/або SSM, і/або LPSM, і/або метаданих будь-якого іншого типу), що ідентифікуються базовим елементом, і захисні величини для щонайменше одного типу метаданих, що ідентифікується за допомогою заголовку сегмента метаданих (або іншими базовими елементами цього сегмента метаданих). Корисне завантаження (декілька завантажень) цього сегмента метаданих іде за заголовком сегмента метаданих, і (у деяких випадках) воно є включеним в базові елементи цього сегмента метаданих.

30 Варіанти здійснення справжнього винаходу можуть бути виконані в апаратному забезпеченні, апаратно-програмному забезпеченні або в програмному забезпеченні, або в їхній комбінації (наприклад, як програмувальна логічна матриця). Якщо не обумовлене інше, алгоритми або процеси, включені як частина винаходу, у своїй внутрішній сутності не відносяться до якого-небудь конкретного комп'ютера або іншого пристрою. Зокрема, із програмами, написаними відповідно до методик цього документа, можна використовувати різні машини загального призначення, або для виконання необхідних етапів способів може бути більш зручно сконструювати більш спеціалізовані пристрої (наприклад, інтегральні мікросхеми).

45 Так, винахід може бути реалізовано в одній або декількох комп'ютерних програмах, що виконуються на одній або декількох програмувальних комп'ютерних системах (наприклад, в одній з реалізацій кожного з елементів за фіг. 1, або в кодері 100 за фіг. 2 (або в одному з його елементів), або в декодері 200 за фіг. 3 (або в одному з його елементів), або в постпроцесорі 300 за фіг. 3 (або в одному з його елементів)), кожна з яких містить щонайменше один процесор, щонайменше одну систему зберігання даних (утримуючу енергозалежну й енергонезалежну пам'ять і/або запам'ятовувальні елементи), щонайменше один пристрій або порт введення й щонайменше один пристрій або порт виводу. Керуючу програму застосовують до вхідних даних для виконання функцій, описаних у цьому документі, і генерування вихідної інформації. Цю вихідну інформацію відомим образом застосовують до одного або декількох пристроїв виводу.

55 Для сполучення з комп'ютерною системою кожна така програма може бути виконана будь-якою належною комп'ютерною мовою (у тому числі машинною, асемблерною або високорівневою процедурною, логічною або об'єктно-орієнтованою мовами програмування). У кожному разі, це мова може являти собою компільовану або інтерпретовану мовою.

60 Наприклад, при реалізації за допомогою послідовностей команд комп'ютерного програмного

забезпечення різні функції й етапи варіантів здійснення винаходу можуть бути реалізовані багатопотоковими послідовностями команд програмного забезпечення, що запускаються на підходящому апаратному забезпеченні цифрової обробки сигналів, і, у цьому випадку, різні пристрої, етапи й функції варіантів здійснення можуть відповідати частинам команд програмного забезпечення.

Кожну таку комп'ютерну програму переважно зберігають у пам'яті або завантажують на носій, або пристрій зберігання даних (наприклад, на твердотільний запам'ятовувальний пристрій, або магнітний або оптичний носій), зчитуване програмувальним комп'ютером загального або спеціального призначення для конфігурування й приведення в дію комп'ютера, коли носій даних або пристрій зчитується комп'ютерною системою для виконання процедур, описаних у цьому описі. Система за винаходом також може бути реалізована як машинозчитуваний носій даних, сконфігурований комп'ютерною програмою (наприклад, що зберігає її в пам'яті), при цьому сконфігурований у такий спосіб носій даних викликає дію комп'ютерної системи специфічним і попередньо визначеним чином для виконання функцій, описаних у цьому документі.

Були описані чисельні варіанти здійснення винаходу. Проте, слід зрозуміти, що не відступаючи від суті й обсягу винаходу можливим є здійснення різних модифікацій. У світлі вищеописаних ідей можливі чисельні модифікації й зміни справжнього винаходу. Слід розуміти, що в межах обсягу формули винаходу, що додається, винахід може застосовуватися на практиці інакше, в порівнянні з описанням конкретно в цьому документі.

ФОРМУЛА ВИНАХОДУ

1. Модуль обробки аудіоданих, що містить:

буферну пам'ять; і

щонайменше одну підсистему обробки даних, пов'язану із зазначеною буферною пам'яттю, при цьому буферна пам'ять зберігає щонайменше один кадр кодованого бітового аудіопотоку, і зазначений кадр містить метадані відомостей про програму або метадані структури вкладених потоків щонайменше в одному сегменті метаданих щонайменше одного поля ігнорованих даних цього кадра й аудіодані щонайменше в ще одному сегменті цього кадра, при цьому зазначена підсистема обробки даних зв'язана й виконана з можливістю виконання щонайменше однієї з наступних операцій: генерування бітового потоку, декодування бітового потоку або адаптивної обробки аудіоданих бітового потоку з використанням метаданих цього бітового потоку, або щонайменше однієї з наступних операцій: аутентифікації, перевірки вірогідності щонайменше аудіоданих або метаданих бітового потоку з використанням метаданих цього бітового потоку, при цьому зазначений сегмент метаданих містить щонайменше одне корисне навантаження метаданих, при цьому зазначене корисне навантаження метаданих містить:

заголовок; і,

після заголовка - щонайменше деякі з метаданих відомостей про програму або щонайменше деякі з метаданих структури вкладених потоків.

2. Модуль обробки аудіоданих за п. 1, у якому зазначений кодований бітовий аудіопотік служить ознакою щонайменше однієї звукової програми, і сегмент метаданих містить корисне навантаження метаданих відомостей про програму, при цьому зазначене корисне навантаження метаданих відомостей про програму

містить:

заголовок метаданих відомостей про програму; і

після заголовка метаданих відомостей про програму - метадані відомостей про програму, що служать ознакою щонайменше однієї властивості або характеристики, звукового вмісту цієї програми, при цьому зазначені метадані відомостей про програму містять метадані активних каналів, що служать ознакою кожного незаглушеного каналу програми й кожного заглушеного каналу програми.

3. Модуль обробки аудіоданих за п. 2, у якому метадані відомостей про програму також містять щонайменше одні з наступних метаданих:

метадані стану обробки понижувального міксування, що служать ознакою того, чи зазнала програма понижувального міксування, і, якщо це так - типу понижувального міксування, яке було застосовано до програми;

метадані стану підвищувального міксування, що служать ознакою того, чи зазнала програма підвищувального міксування, і, якщо це так - типу підвищувального міксування, яке було застосовано до програми;

метадані стану попередньої обробки даних, що служать ознакою того, чи виконувалася на звуковому вмісті цього кадра попередня обробка даних, і, якщо це так - типу попередньої обробки даних, виконаної на зазначеному звуковому вмісті; або

5 метадані обробки розтягування спектра або зв'язування каналів, що служать ознакою того, чи застосовувалася до програми обробка розтягування спектра або зв'язування каналів, і, якщо це так - діапазону частот, до якого було застосоване розтягування спектра або зв'язування каналів.

4. Модуль обробки аудіоданих за п. 1, у якому зазначений кодований бітовий аудіопотік служить ознакою щонайменше однієї звукової програми, що містить щонайменше один незалежний вкладений потік звукового вмісту, і сегмент метаданих містить корисне навантаження метаданих структури вкладених потоків, при цьому зазначені метадані структури вкладених потоків містять:

заголовок метаданих структури вкладених потоків; і

після заголовка метаданих структури вкладених потоків - метадані незалежного вкладеного потоку, що служать ознакою кількості незалежних вкладених потоків програми, і метадані залежних вкладених потоків, що служать ознакою того, чи містить кожний незалежний вкладений потік програми щонайменше один асоційований залежний вкладений потік.

5. Модуль обробки аудіоданих за п. 1, у якому сегмент метаданих містить:

заголовок сегмента метаданих;

після заголовка сегмента метаданих - щонайменше одну захисну величину, придатну для щонайменше однієї з операцій: розшифрування, аутентифікації або перевірки вірогідності щонайменше одних з наступних даних: метаданих відомостей про програму, метаданих структури вкладених потоків або аудіоданих, відповідних до зазначених метаданих відомостей про програму або зазначених метаданих структури вкладених потоків; і

після заголовка сегмента метаданих - ідентифікатор корисного навантаження метаданих і конфігураційні значення корисного навантаження, при цьому корисне навантаження метаданих іде за ідентифікатором корисного навантаження метаданих і за конфігураційними значеннями корисного навантаження.

6. Модуль обробки аудіоданих за п. 5, у якому заголовок сегмента метаданих містить синхрослово, що ідентифікує початок сегмента метаданих, і за синхрословом іде щонайменше одна величина-ідентифікатор, і заголовком корисного навантаження метаданих містить щонайменше одну величину-ідентифікатор.

7. Модуль обробки аудіоданих за п. 1, у якому зазначений кодований бітовий аудіопотік являє собою бітовий потік AC-3 або бітовий потік E-AC-3.

8. Модуль обробки аудіоданих за п. 1, у якому зазначена буферна пам'ять зберігає кадр неминущим чином.

9. Модуль обробки аудіоданих за п. 1, при цьому цей модуль обробки аудіоданих являє собою кодер.

10. Модуль обробки аудіоданих за п. 9, у якому зазначена підсистема обробки даних містить:

підсистему декодування, виконану з можливістю приймання вхідного бітового аудіопотоку й добування вхідних метаданих і вхідних аудіоданих із цього вхідного бітового аудіопотоку;

підсистему адаптивної обробки даних, зв'язану й виконану з можливістю виконання адаптивної обробки даних на вхідних аудіоданих з використанням вхідних метаданих, за допомогою чого генеруються оброблені аудіодані; і

підсистему кодування, зв'язану й виконану з можливістю генерування кодованого бітового аудіопотоку у відповідь на оброблені аудіодані, що полягає у включенні метаданих відомостей про програму або метадані структури вкладених потоків у зазначений кодований бітовий аудіопотік і в напрямку цього кодованого бітового аудіопотоку в буферну пам'ять.

11. Модуль обробки аудіоданих за п. 1, при цьому цей модуль обробки аудіоданих являє собою декодер.

12. Модуль обробки аудіоданих за п. 11, у якому зазначена підсистема обробки даних являє собою підсистему декодування, пов'язану з буферною пам'яттю й виконану з можливістю добування метаданих відомостей про програму або метаданих структури вкладених потоків із зазначеного кодованого бітового аудіопотоку.

13. Модуль обробки аудіоданих за п. 1, що містить:

підсистему, пов'язану з буферною пам'яттю й виконану з можливістю добування метаданих відомостей про програму або метадані структури вкладених потоків із зазначеного кодованого бітового аудіопотоку й для добування аудіоданих із цього кодованого бітового аудіопотоку; і

постпроцесор, пов'язаний із цією підсистемою й виконаний з можливістю виконання адаптивної обробки даних на зазначених аудіоданих з використанням щонайменше метаданих одного з

типів: метаданих відомостей про програму або метаданих структури вкладених потоків, витягнутих із зазначеного кодованого бітового потоку.

14. Модуль обробки аудіоданих за п. 1, при цьому цей модуль обробки аудіоданих являє собою процесор цифрової обробки сигналів.

5 15. Модуль обробки аудіоданих за п. 1, при цьому цей модуль обробки аудіоданих являє собою препроцесор, виконаний з можливістю добування метаданих відомостей про програму або метаданих структури вкладених потоків і аудіоданих з кодованого бітового аудіопотоку й для виконання адаптивної обробки даних на цих аудіоданих з використанням щонайменше метаданих одного з типів: метаданих відомостей про програму або метаданих структури
10 вкладених потоків, витягнутих із зазначеного бітового аудіопотоку.

16. Спосіб декодування кодованого бітового аудіопотоку, при цьому зазначений спосіб включає етапи:

приймання кодованого бітового аудіопотоку; і

15 добування метаданих і аудіоданих із цього кодованого бітового аудіопотоку, при цьому зазначені метадані являють собою або містять метадані відомостей про програму й метадані структури вкладених потоків,

при цьому зазначений кодований бітовий аудіопотік містить послідовність кадрів, що служать ознаками щонайменше однієї звукової програми, зазначені метадані відомостей про програму й метадані структури вкладених потоків служать ознаками зазначеної програми, кожний з кадрів
20 містить щонайменше один сегмент аудіоданих, кожний зазначений сегмент аудіоданих містить щонайменше деякі з аудіоданих, кожний кадр із щонайменше підмножини кадрів містить сегмент метаданих, і кожний зазначений сегмент метаданих містить щонайменше деякі з метаданих відомостей про програму й щонайменше деякі з метаданих структури вкладених потоків.

25 17. Спосіб за п. 16, у якому зазначений сегмент метаданих містить корисне навантаження метаданих відомостей про програму, при цьому зазначене корисне навантаження метаданих відомостей про програму містить:

заголовок метаданих відомостей про програму; і

після заголовка метаданих відомостей про програму - метадані відомостей про програму, що
30 служать ознакою щонайменше однієї властивості, або характеристики, звукового вмісту цієї програми, при цьому зазначені метадані відомостей про програму містять метадані активних каналів, що служать ознакою кожного незаглушеного каналу програми й кожного заглушеного каналу програми.

18. Спосіб за п. 17, у якому метадані відомостей про програму також містять щонайменше одні з наступних метаданих:

35 метадані стану обробки понижувального мікшування, що служать ознакою того, чи зазнала програма понижувального мікшування, і, якщо це так - типу понижувального мікшування, яке було застосовано до програми;

40 метадані стану підвищувального мікшування, що служать ознакою того, чи зазнала програма підвищувального мікшування, і, якщо це так - типу підвищувального понижувального мікшування, яке було застосовано до програми; або

метадані стану попередньої обробки даних, що служать ознакою того, чи виконувалася на звуковому вмісті цього кадра попередня обробка даних, і, якщо це так - типу попередньої обробки даних, виконаної на зазначеному звуковому вмісті.

45 19. Спосіб за п. 16, у якому зазначений кодований бітовий аудіопотік служить ознакою щонайменше однієї звукової програми, що містить щонайменше один незалежний вкладений потік звукового вмісту, і сегмент метаданих містить корисне навантаження метаданих структури вкладених потоків, при цьому зазначені метадані структури вкладених потоків містять:

заголовок метаданих структури вкладених потоків; і,

50 після заголовка метаданих структури вкладених потоків - метадані незалежного вкладеного потоку, що служать ознакою кількості незалежних вкладених потоків програми, і метадані залежних вкладених потоків, що служать ознакою того, чи містить кожний незалежний вкладений потік програми щонайменше один асоційований залежний вкладений потік.

20. Спосіб за п. 16, у якому зазначений сегмент метаданих містить:

55 заголовок сегмента метаданих;

після заголовка сегмента метаданих - щонайменше одну захисну величину, придатну щонайменше для однієї з операцій: розшифрування, аутентифікації або перевірки вірогідності щонайменше одних з даних: метаданих відомостей про програму, метаданих структури вкладених потоків або аудіоданих, відповідних до зазначених метаданих відомостей про
60 програму або зазначених метаданих структури вкладених потоків; і

після заголовка сегмента метаданих - корисні навантаження метаданих, що містять щонайменше деякі з метаданих відомостей про програму або щонайменше деякі з метаданих структури вкладених потоків.

21. Спосіб за п. 16, у якому зазначений кодований бітовий аудіопотік являє собою бітовий потік AC-3 або бітовий потік E-AC-3.

22. Спосіб за п. 16, що також включає етап: виконання адаптивної обробки на аудіоданих з використанням щонайменше метаданих одного з типів: метаданих відомостей про програму або метаданих структури вкладених потоків, витягнутих із зазначеного кодованого бітового потоку.



Fig. 1

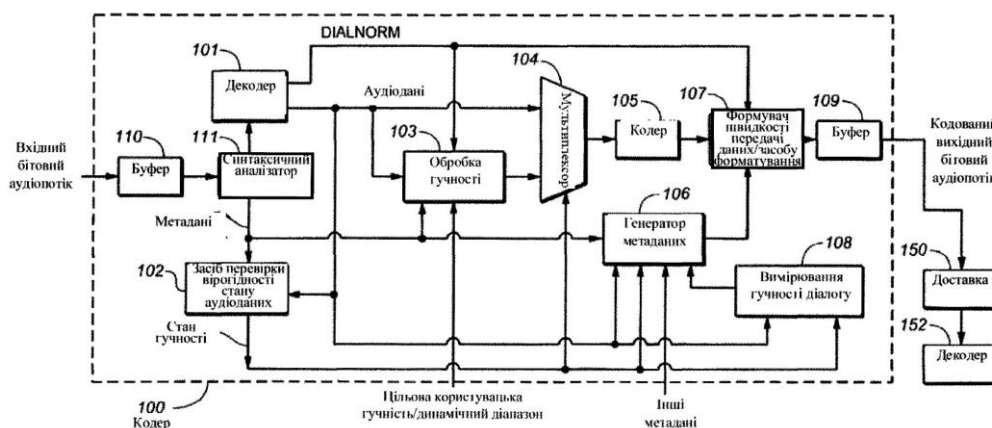


Fig. 2

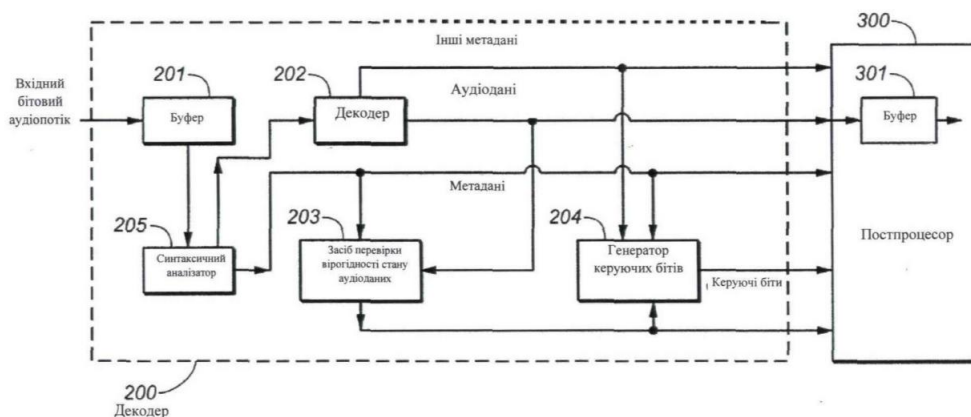


Fig. 3

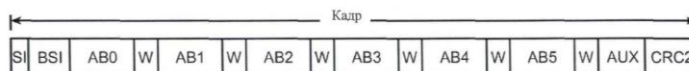


Fig. 4

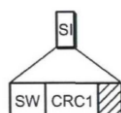


Fig. 5



Fig. 6

SI	BSI	one to six of AB0, AB1, AB2, AB3, AB4, AB5	W	AUX	CRC2
----	-----	--	---	-----	------

Фіг.7
Контейнер



(Структура контейнера (сегмента метаданих))

Fig.8

Комп'ютерна верстка А. Крулевський

Державна служба інтелектуальної власності України, вул. Василя Липківського, 45, м. Київ, МСП, 03680, Україна

ДП “Український інститут інтелектуальної власності”, вул. Глазунова, 1, м. Київ – 42, 01601